

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 27, 2001, 14:35:56 ; Search time 3871.54 Seconds
(without alignments)
11781.976 Million cell updates/sec

Title: US-09-541-094-13

Perfect score: 2949
Sequence: 1 tctgcagaatcgcgtctgcg.....attaataatagatatccacc 2949

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rud:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rud:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_r01:*
95: gb_r02:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2948.2	100.0	2949	89 AF240468	AF240468 Homo sapi
2	2804.2	95.1	2805	91 D87442	D87442 Human mRNA
3	1746.4	59.2	2250	94 AF240469	AF240469 Mus muscu
4	802	27.2	155913	81 AL513282	AL513282 Homo sapi
5	802	27.2	182344	81 AL445230	AL445230 Homo sapi
6	791	26.8	170677	78 AL138932	AL138932 Homo sapi
7	332	11.3	165942	91 AP001605	AP001605 Homo sapi
8	332	11.3	340000	91 AP001699	AP001699 Homo sapi

C 9	325.6	11.0	150597	77	AC090058	AC090058 Homo sapi
C 10	262.6	8.9	223431	75	AC074310	AC074310 Mus muscu
C 11	256.2	8.7	190442	66	AC021119	AC021119 Homo sapi
C 12	213.6	7.2	170677	78	AL138932	AL138932 Homo sapi
C 13	180.8	6.1	148540	81	AL391417	AL391417 Homo sapi
C 14	180.8	6.1	167253	90	AL445926	AL445926 Human DNA
C 15	180.8	6.1	168537	82	AL590397	AL590397 Homo sapi
C 16	180.8	6.1	173152	70	AC026466	AC026466 Homo sapi
C 17	180.8	6.1	246646	81	AL391275	AL391275 Homo sapi
C 18	180.8	6.1	302665	81	AL391377	AL391377 Homo sapi
C 19	144.6	4.9	4692	97	HS024105	HS024105 Homo sapi
C 20	79.2	2.7	7218	10	166494	166494 Sequence 14
C 21	59.4	2.0	8258	53	CNS060YX	AL416479 T7 end of
C 22	54.4	1.8	2909	5	AF240470	AF240470 Drosophila
C 23	54	1.8	187540	75	AC073970	AC073970 Homo sapi
C 24	53.2	1.8	154608	74	AC073620	AC073620 Homo sapi
C 25	51.4	1.7	72500	77	AC090206	AC090206 Homo sapi
C 26	50.6	1.7	258174	76	AC079429	AC079429 Mus muscu
C 27	50	1.7	101509	70	AC027353	AC027353 Homo sapi
C 28	50	1.7	148600	70	AC027275	AC027275 Homo sapi
C 29	49.8	1.7	53522	97	H0MPKD1GEN	L39891 Homo sapien
C 30	49.8	1.7	53526	9	AR097326	AR097326 Sequence
C 31	49.8	1.7	53577	9	AR097325	AR097325 Sequence
C 32	49.4	1.7	190492	75	AC073957	AC073957 Homo sapi
C 33	49.4	1.7	298166	77	AC087563	AC087563 Homo sapi
C 34	49	1.7	1063	53	CNS0616W	AL403806 T3 end of
C 35	48.4	1.6	58998	63	AC014104	AC014104 Drosophila
C 36	48.4	1.6	215711	5	AC008205	AC008205 Drosophila
C 37	48.4	1.6	228479	5	AE003749	AE003749 Drosophila
C 38	48.2	1.6	139524	78	AC090444	AC090444 Rattus no
C 39	47.8	1.6	5809	85	AB007974	AB007974 Homo sapi
C 40	47.8	1.6	92775	90	AL441964	AL441964 Human DNA
C 41	47.8	1.6	213432	73	AC068198	AC068198 Homo sapi
C 42	47.8	1.6	267759	76	AC079418	AC079418 Mus muscu
C 43	47.6	1.6	14955	5	AF000580	AF000580 Dictyoste
C 44	47.6	1.6	152513	83	AP003207	AP003207 Oryza sat
C 45	47.6	1.6	165608	77	AC087842	AC087842 Rattus no

ALIGNMENTS

RESULT	1	2949 bp	mRNA	PRI	08-SEP-2000
LOCUS	AF240468				
DEFINITION	Homo sapiens nicastrin mRNA, complete cds.				
ACCESSION	AF240468				
VERSION	AF240468.1	GI:9992877			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				

REFERENCE 1 (bases 1 to 2949)
 Yu, G., Nishimura, M., Arawaka, S., Levitan, D., Zhang, L., Tandon, A., Song, Y.-Q., Rogava, E., Chen, F., Kawarai, T., Supala, A., Levesque, L., Yu, H., Yang, D.-S., Holmes, E., Milman, P., Liang, Y., Zhang, D. M., Xu, D. H., Sato, C., Rogava, E., Smith, M., Janus, C., Zhang, Y., Abersold, R., Farrer, L. S., Sorbi, S., Bruni, A., Fraser, P. and St George-Hyslop, P.

TITLE
 Nicastrin modulates presenilin-mediated notch/gip-1 signal transduction and betaAPP processing
 Nature 407 (6800), 48-54 (2000)

JOURNAL MEDLINE
 20045163

REFERENCE
 2 (bases 1 to 2949)
 Yu, G., Nishimura, M., Arawaka, S., Levitan, D., Zhang, L., Tandon, A., Song, Y.-Q., Rogava, E., Kawarai, T., Supala, A., Abersold, R., Sorbi, S., Bruni, A., Fraser, P. E. and St George-Hyslop, P. H.

TITLE
 Submitted (02-MAR-2000) Dept. of Medicine, Centre for Research in Neurodegenerative Diseases, University of Toronto and University Health Network, Rm. 118, Tanz Neuroscience Bldg., 6 Queen's Park Crescent West, Toronto, Ontario M5S 3H2, Canada

FEATURES	Location/Qualifiers
source	1..2949
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="near D12S1595 and DIS2844"
CDS	143..2272
	/function="modulates beta-APP processing and Notch/gip1 signal transduction"
	/note="type 1 transmembrane protein; similar to Mus musculus and Drosophila melanogaster nicastrin encoded by GenBank Accession Numbers AF240469 and AF240470; interacts with presenilin proteins PS1, PS2, S182, and E5-1"
	/codon_start=1
	/product="nicastrin"
	/protein_id="AA011412.1"
	/db_xref="GI:9992878"
	/translation="MATAGSGGADPGSGNGLRLSPVLAGLCRGSVKRIYPLNTPAPVRLNATHOIGCSSISGTCVYHYEKEDLQWVLDGPNPVLLESKHETRDMEKLGKRTSRIGLAVSLTPSPASGSPVQC.PNDGFVYSNSYPERACHRTIOWNSLGNGLAVEDFSPFIFLEDENETKVKCYODHNI.SONGSAPFPLCAMOLF SHMAVISTATCMRRSSIOSTFISINPEIYCDPLSDYNWSMLKPIINTGTLKPDROY VATRIDSRFPMVYAPGASVAFVYQLAAEALOKAPVTTILPRMVFPOGERT FVIGSSRMVYDMEKGPVQLENDSPVEIYGOVALFSLIEMHTDPVSOQNSVRN KYOSTIDPRAENINYSYPMISPEEDLFTYDTRAKLADVATFVIGRALYELAGTNS DYQADPQVITLTLGLFKANNSWFSILKQDLKSYIGDPLQHYLAVSSPTWTYV VYALANLGTIVNLTRQCDOPSPVSENLDEYSWVQPLHSNEDRLRCVRSF ABLARLSAPFEL.SOMSTSEYESTWSEMKDIRAIFLASKELTLITLVFGFILLF SLIVRSLINAKADVLFIAIPREGAASY"
BASE COUNT	682 a 803 c 692 g 770 t 2 others
ORIGIN	
Query Match	100.0%; Score 2948.2; DB 89; Length 2949;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2949; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 tctcgaattcggcttgccttcgaaacgaacttcggcttccttggtccggcgca 60
DB	1 TCTGCAAGTAATTCGGCTTTCGCGCTCGGAACACGACTTCGGCTCTTAGCGTCCGGGCA 60
QY	61 cagaacggtgtcgaatgtgtagccttagagagccgctaaacagacgagacgagggcg 120
DB	61 CAGAGACGGTGTCTAGGTGAGTACGCTAGAGAGCGCGCTAACACAGAGACCGACGGGCG 120
QY	121 ttccgctcagagagagagaaatgctacgagcaggggtggtcttgggtggtaccggg 180
DB	121 TTCCGCTCAGCAGAGAGGGAAGATGATATATCCCTTAATAAACAGCTCCCTGTTCCGCT 180
QY	181 aagtcggggtctccttcgctctcttcttcgtcgtcctactacgagttgttcagagg 240
DB	181 AAGTCGGGGTCTCTTCGCGCTCTCTTTCGCGCTCTTACTAGACAGTTGTGTGACGGG 240
QY	241 aaactcagtgagagagagatatataatcccttaataaacaagctccgtgttcgct 300
DB	241 AAACCTCAGTGAGAGAGAGATATATATCCCTTAATAAACAGCTCCCTGTTCCGCT 300
QY	301 gttcaacgcactcactcagatggtggtcgcagctctcaattgttgagaacagggttat 360
DB	301 GGTCAACGCCACTCTATCAGATTTGCTGCCAGTCTTCAATATGTGGAGACACAGGGTTAT 360
QY	361 ccaagtagtagaagaagaagaagactacagtggtatgtagatgagcccaacccccc 420
DB	361 CCACGTAGTAGAAGAAGAAGAGAGACCTACAGTGGGTATGTACATGCCCCAACCCCCC 420
QY	421 ttacatggttcgtcgtgaggaacagcatattcaacagggatttaaggaagctgaagg 480
DB	421 TTACATGTTCTGCTGGAGACCAAGCATTTTACCAAGGATTTAATGAGAGAAGCTGAAGG 480
QY	481 gagaacagccgaattgtgtcttgacagtgtccttaccagacccagcttcctcagg 540
DB	481 GAGAACCAAGCCGANTGTGCTTTCGACAGTCTCTTACCAAGCCCAAGTCTGCTCAGG 540

QY 541 ctctctctctagtgtagagtgcccaaatgtagtggttggttgtagttagtcccaattccctatgg 600
DB 541 CTCTCTCTCTAGTGAGAGTGCCCAAAATGATGGTTGGTTGTTACTTCACATTCCTATAGG 600
QY 601 gccagagtttgcctacgtgcagagaataatcaatgtagtgcgttgtaggaattgtttggtcta 660
DB 601 GCCAGAGTTTGCCTACGTCAGAGAAATACAGTGTGAGTTGGTGGCAATGGTTGGCTTA 660
QY 661 tgaagaacttgattccccactcttctctcttgaaatgaaatgaaacaaagtcacaa 720
DB 661 TGAAGAATTTGATTCCCCACTCTTCTCTTGAGATGAAATGAACCAAGTCATCAA 720
QY 721 gcaagtgcatacaagaatcaaacactgagtagaattggtctcaagcaaacacttccactatg 780
DB 721 GCAAGTGCATCAAGATCAACACCTGAGTACAGATGGCTCAGCACCAACCTTCCCACTATG 780
QY 781 tgcagatgcagctcttctcaacatgtagtgcgtgtagtcaagcaactgcacactgtagcgag 840
DB 781 TGCAGATGCAGCTCTTTCACACATGATGCTGATCAGCACTGCACCTGACATGGCGG 840
QY 841 gacgtccatccaaagcaccttcaagcatcaacccagaatctgtctgtgacccctgtctga 900
DB 841 GAGCTCATCCAAAGCACCTTACGATCAACCCAGAAATGCTGTGACCCCTGTCTGA 900
QY 901 ttacaatgtgtgtagcatgtagtcaagcctaataatacaactggagacataaagcctgacga 960
DB 901 TTACAATGTTGTGAGCATGCTTAAAGCCTTAAATACAACTGGAGACATTAAAGCCTGACGA 960
QY 961 cagggttgtgtgtgtgtgcacacccggtgtgtagtgcgttcccttcttggaattgtggccc 1020
DB 961 CAGGGTTGTGTTGTGTGTCACCCGGGTGGATAGTGTTCCTTTCGTGAATGTGGCCCC 1020
QY 1021 aggggtgtagaagcagtagtgccttctctgtcaacccagctgtagtgcgtgtagaacttgca 1080
DB 1021 AGGGGTGAGAAAGCAGTGTGCTTCTTGTCCACCAAGCTGCTGCTGGAAGCTTTGCA 1080
QY 1081 aaaggaacctgtagtgcacacccctgcaccccgaaatgtagtgcgttctctcttcaaggaga 1140
DB 1081 AAAGGACCTGAGTGTGACACCCCTGCCCCGCAATGTCAATGTTCCTTTCCTTCAAGGGGA 1140
QY 1141 aactttagctacatgttagagagctgcagagtagtctacagataagagagaggaacttcc 1200
DB 1141 AACTTTGACTACATTTGGCAGCTCCAGGATGTCTTACGATATGGAGAGGCAAGTTTCC 1200
QY 1201 cgtgcagttagaagaatgttgaactcatctgtgagctgtaggacagtgtagccttaagaacttc 1260
DB 1201 CGTGCAGTTAGAGAAATGTTGACTCATTTGTGGAGCTGGAGAGGTGGCCTTAAAGAACTTC 1260
QY 1261 attagaagcttggatgacacagatccgttctcagaataatgtagtgcgtgtagaagca 1320
DB 1261 ATTAGAGCTTTGGATGTCACACAGATCTCTTCTCAGAAATATGATGTGTGAGGAAACA 1320
QY 1321 ggtgagagagcttccctgcagacatgtgagaagagtggtgtgtgtgtgtgtgtgtgtgtgt 1380
DB 1321 GGTGAGAGATCTCTGCTGCGACATGTGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
QY 1381 caagagagcaaatcaagctccagcctctccacacatcttccctgcagagatcttccctagagc 1440
DB 1381 CAGAGAGCCAAATCAATCCAGCCTTCCACCATCTTCCCTGACAGATTTCTTTCGAGAC 1440
QY 1441 tgcgaacatctcgtggt 1500
DB 1441 TCGAAACATCTCTGCGT 1500
QY 1501 ccagagattttagacacagctgtagaacaatlaattgtagctatcccgaaatggctgagccc 1560
DB 1501 CCAGAGATTTTAGACACAGCTGTGAGAAATTAATGTGAGCTATCCCGAATGGCTGAGCCC 1560
QY 1561 tgaagagagctgtagaacttgttacaagacactgcgaagccctggagagatgtgtgccaaggt 1620
DB 1561 TGAAGAGAGCTGAGACTTTGTAAACAGACACTGCCAAGGCCCTGGAGATGTGGCCACGGT 1620

QY 1621 gctggagcgtgctctgtatgagcttgcaaggaggaacaaacttccagcagtagtccagc 1680
DB 1621 GCTGGAGCTGCTCTGTATGAGCTTCCAGAGGAGAACCAACTTCAGGACACTTAGGC 1680
QY 1681 tgaatcccaagcgttaacccgctgtcctatgagtggttctgtatataaagccaactcatg 1740
DB 1681 TGAATCCCAAGCGTTTACCCGCTGTGCTATGGGTTCTGTATTAAGGCCAACAACTCATG 1740
QY 1741 gtccagctatccctcaagcagaactaagttccacttgggtgtagcggccttcaaca 1800
DB 1741 GTTCCAGCTATCTCTCAGCAGGACCTTAAAGTCTTCTGAGGAGCGGCTTCTTAAACA 1800
QY 1801 ttaaatgcgtcttccagccccaacacacttattgttgaactatgagcttggcaaa 1860
DB 1801 TTAAATGCTGTCTCCAGCCCCAACCAACCACTTATGTGTAGATAGCTTGGCAAA 1860
QY 1861 tttagctgcaagtagtcaacactcaacccagagagtagtgcagaatcaagtaagttccc 1920
DB 1861 TTTGACTGGCAGAGTGTCAACCTCAACCGAGAGAGATGCCAGAGATCAAGTAAAGTCCC 1920
QY 1921 aagtgaataaagaatctgtatgtagtcaatggtgtccagggccttgcataatga 1980
DB 1921 AAGTGAATAAAGAGATCTGTATGAGTACTCATGGGTCCAGGGCCCTTGTGATTCATGA 1980
QY 1981 gacggaccgactcccccgggt 2040
DB 1981 GACGGACCGACTCCCGGGT 2040
QY 2041 tgccttgaactgtagtgcagtagtgccttcaatgatacttcaatgtagtgcagtagcgtg 2100
DB 2041 TGCTTTGAATGTAGTCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 2100
QY 2101 gaaagataatccgtgcagagatatttccatccagcagaagagctgtgtagtgcacct 2160
DB 2101 GAAAGATATCCGTCCCGGATATTTCATGCGCAGCAAGAGCTGTGATGTGATCACCT 2160
QY 2161 gacagtgagcttgtagtgccttccatcttccctcatcgttcaactgtagtgcagtagc 2220
DB 2161 GACAGTGGGCTTCGGCATCTCATCTTCTCCATCATGTGTCACCTACGATTCATGCAATGCA 2220
QY 2221 agctgaagcttcttcaatgtctcccgaggagcagagctgtgtagtgcagtagcagtagc 2280
DB 2221 AGCTGATGTCTTTCATATGCTCCCGGAGCCAGGAGCTGTGCAATACAGASAGAGCC 2280
QY 2281 scaagcttcttgcagctgcagctgcagctgcagctgcagctgcagctgcagctgcagctgc 2340
DB 2281 SCAGCTTTTCTTGGCCAGCTCAGCAAGTTCATCTTCAAGAGATGTGTCCACCTGGGACACA 2340
QY 2341 accactaatgtgcactggaacccctccctgagcctgctcagatggatlaacataaag 2400
DB 2341 ACCACTAATTTGTGCACTGGAACCTCCGCGGCTGTCTCAATTTGGGATTAACATAAAG 2400
QY 2401 agtgaactatccaaagaagaagaggaataaataatgtcctccctccctccctcc 2460
DB 2401 AGTGAACATATCCAAAGAGAGAGAGGAGAAATTAATTAATGCTCCCTCTCCCTCCCTCC 2460
QY 2461 ccttcccatcaacccctcccatcttctctctctctctctctctctctctctctctctct 2520
DB 2461 CTTTCCCATACACCTTCTCCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
QY 2521 ttacaatagaagctctctgtcctgtttaaactccctagttaccacactaattgcct 2580
DB 2521 TTACAATAGAAGCTTCTGTGCTGCTTAACTCCTAGTACCACACCTTAATTTGGCCT 2580
QY 2581 tcaaggaaccttcaatttcttccctccctgtgtagtgccttctgtccctccacccagc 2640
DB 2581 TCAGGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2640
QY 2641 cctgtaccagacacttccctgtagtgcagtaggagacataaaggtttaaattgtaggttcaaa 2700
DB 2641 CCTGTACCAGACACTTCTCTGACTGTGAGAGAGACATAAAGGTTTAATGTACAGGTCAAA 2700
QY 2701 ctacattgagcccttgaggaagagagtagtctctggtctgtagcctactgctctctccac 2760

Query Match	Best Local Similarity	95.1%;	Score 2804.2;	DB 91;	Length 2805;			
Matches 2803;	Conservative	2;	Mismatches	0;	Indels	0;	Gaps	0;
QY	145	ggtacagcaggggtgtgctctcgtggagctgacccgggaagtcggggctctctcgcctct	204					
Db	1	GGCTACGGCAGGGGGTGGCTCTGGGGCTGACCCGGGAAGTGGGGTCTCTCGCTTCT	60					
QY	205	ggtcttcgtcgctcactagcggcttctgtcaggggaaacctcagtggaggaagata	264					
Db	61	GTCTTCTCGCTCCTCACTAGCAGGTTTGTGCAAGGGGAAATCAGAGGAGAGAGATATA	120					
QY	265	tatcccttaataaacaacgctccctgtgttcgctgtctcaacgccaatcaagattgg	324					
Db	121	TATCCCTTAATAAACAAGCTCCTGTGTTGGCTGTGCTCAACGCCATCTATCAGATTGG	180					
QY	325	ctgcagatcttcaattagtgtgagacacaggggtatccaagttagtagaagaaggaga	384					
Db	181	GTGCCAGTCTTAAATTAATGAGGACACAGGGGTTATCCACGTAGTAGAGAAAGGAGGA	240					
QY	385	ccctacagtggtattgcactgacgtatgcccccaacccccctacaigtcttcgtcggagaa	444					
Db	241	CCTACAGTGGGATTGACTGATGAGGCCCAACCCCTTACATGGTTCTGCTGGAGAGCAA	300					
QY	445	gcattttaccaggaattaaatgaggaagctgaaagggaacacagccgaattgtgtct	504					
Db	301	GCATTTTACCAAGGATTTAATGGAACACTGAAGGGAGAACACAGCAATTGCTGCTGCT	360					
QY	505	tgcagttgccttgaccacaagccagtcgcgtctcgttcaggtctctcctagtgaacgtccc	564					
Db	361	TGCAGTGTCTTGTACCAAGCCCAAGTCTCGCTCAAGGCTTCTCTCTATGATCAGTCCCT	420					
QY	565	aaatgattggttgttgttactccaattccatgagggcagagtttgtcactcagaga	624					
Db	421	AAATGATGGGTTTGTGTACTCCAAATTCCTATGAGGCGAAGTTTCTCTACTGCAAGAA	480					
QY	625	aatacagtggaattcgcctggtggcaatggtttggtctatagaagactttagttcccactct	684					
Db	481	AATACAGTGGAATTCGCTGGGCAATGTTTGGCTTATGAACACTTGTAGTTTCCCATCTT	540					
QY	685	tcttcttgaagaatgaaatgaaaccaaggtatcaagcagtggtcatataagaatcaaacct	744					
Db	541	TCTTCTTGAAGATGAATAATGAACCAAGTATATAGCAGTGTCTATATAGTACAAACT	600					
QY	745	gagtcagaatggtctcagcaccaaccttcccactatgtgtccatgtagctcttcttcaacat	804					
Db	601	GAGTCAGATAGGCTCAGACCAACCTTCCCACTATGTGCGCATGTCAGCTTTTCCACAT	660					
QY	805	gcattgtctatagaacttgcacctgtgcatgctgagctccatctcaagaagcaacttcag	864					
Db	661	GCATGCTCTATCAGCACTGACCTCGATGCGGCGGAGCTCCATCCAAAGCACTTTCAG	720					
QY	865	catcaaccagaataatgctcttgacccccctgtcgattacaatgtgtggagatgtctaaa	924					
Db	721	CATCAACCCAAATTCGTCTGTGACCCCTGTCTGATTACATATGTGTGGAGCATGCTAAA	780					
QY	925	gctctataataacaactgggacattaaagctctgacgacagaggttgtgtgtgtccaccg	984					
Db	781	GCTATTAATTAATCACTGGGACATTAAAGCCTGACGACAGGTTGTGTGTTCTCTCCACCG	840					

QY 985 gctgagatgctgcttcttctggaatgtgcccagggctgaaagcagtgcttc 1044
 Db 841 GCTGGATGATGCTCTTCTTCTGGAATGCGCCAGGGGCTGAAAGGCGATGCTTC 900
 QY 1045 cttgtaccagctgctgctgctgaagcttgcgaagcagcctgactgacaccc 1104
 Db 901 CTTTGTACCCAGCTGGCTGCTGCTGAAGCTTGCAAAAGGACCGATGTGACCCCT 960
 QY 1105 gccccgaatgctcattgttcttcttcaaggagaaacttctgactacaatggcagtc 1164
 Db 961 GCCCCGAATGCTCATGTTTCTTCTTCAAGGGGAAACTTTTGACTACATGGCGCTTC 1020
 QY 1165 gagatgtgctagatatagtgaaagagcaggttcccgctgcaatgagaatgtgac 1224
 Db 1021 GAGGATGCTACGATATGGAGAAAGGCAAGTTCCCGTGCATTTGAGAGATTTACTTC 1080
 QY 1225 attgtgagctgagagagtgctgcttcaagaactcattagagcttggatgcacaga 1284
 Db 1081 ATTGTGAGCTGGGACAGGTGGCTTAAGAACTTCAATAGAGCTTGGATGCACACAGA 1140
 QY 1285 tccgttctcagaagaatgtgctgtaagcagagtgagagatctcctgacacatt 1344
 Db 1141 TCTCTTCTCAGAAAAATGAGTGTGACGGAACCAAGTGGAGATCTCTGGCCACATT 1200
 QY 1345 gagaagagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1404
 Db 1201 GGAGAAAGT 1260
 QY 1405 tctccacacatctccctgcaagcatttcttgcagctcgaagctcctgctgctgct 1464
 Db 1261 TCTCCACACATCTCCCTGCAAGCTTCTGAGCTGGAACATCTGTGCTGTGTGTCT 1320
 QY 1465 ggtcagcactgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1524
 Db 1321 GGTGACCACTGT 1380
 QY 1525 gaacatlaatgtgactatcccgaaatgtgctgagcccgaaagagcctgaaattgtaac 1584
 Db 1381 GAACATTAATGTGACTATCCGAAATGTGCTGAGCCCTGMAAGAGACCTGAACCTTTGTAC 1440
 QY 1585 aagacatgcaagcagccctgagcagatgtgagcagctgctgagcagctgctgctg 1644
 Db 1441 AGACACTGCAAGCCCTGCGAGATGTGCGACGCTGTGGAGCTGTGTGTGTGTGTGT 1500
 QY 1645 tgcagagagaaacaaactcaagcagacagctgctgctgctgctgctgctgctgct 1704
 Db 1501 TGCAGAGAGAACCAACTTCAAGCAGACAGTTCAAGCTGATCCCAAGGTTTACCCGCT 1560
 QY 1705 gctctatggttctgactaaagcacaacactcagttcagttcagttcagttcagtt 1764
 Db 1561 GCTCTATGGGTTCTGATTAAGCCAAACATCATGTGTCTCAAGTATCTCTCAAGCAGGA 1620
 QY 1765 cctaaggtcctaactgtgctgagcagcctctcaacatcagcgtgctgctgctgctg 1824
 Db 1621 CCTAAGTCTCTAATTGGGTGAGCGGCTCTTCAACATTCATCGCTGCTCAGCCAC 1680
 QY 1825 caaacacaactatgttgcacagatgctgctgcaaatctgactgctgcaagtggttaact 1884
 Db 1681 CAACACCACTTATGTGTACAGATGCTTGGCAAAATTTGACTGGCACAGTGTAACTT 1740
 QY 1885 caccagagagcagtgccagagatcacaagtaagttcccaagtgaacaagagatctalga 1944
 Db 1741 CACCCGAGACAGTGCAGAGATCCAAGTAAAGTCCCAATGGAAGCAAGAGATCTTATGA 1800
 QY 1945 gtaactcagtgccagggccttctgactcactaagagagcagcagctcccggtgtgt 2004
 Db 1801 GTACTCATAGGCTCAGGGCCCTTGCATTTCTATGAGAGCGACCTCCCGGTGTGT 1860
 QY 2005 gctgttcaatgagcagatatacagggcctgtgtcctgctgcttgaacagtgagtgag 2064
 Db 1861 GCGTTTACTGACAGATATTAAGCCAGGGCCTTGTCTCTGCTTGAACGTAGTACGTGAG 1920

QY 2065 cctactgaatactctacatgagctgagagccgctggaagaatctcgttgcocagat 2124
 Db 1921 CTTACTGATTAATCTTACATGAGACTGAGAGCCGCTGGAAGATATCCGCGGATATT 1980
 QY 2125 tctatgcagaagaagagcttgaatgatacaccctgagcagtgagcttgcagctcat 2184
 Db 1981 TCTATGCGCAGCAAAAGCTTGAATGATGATCACCCTGACAGATGGGCTTGGCATCTCAT 2040
 QY 2185 cttccctcatcgtcactcactgactcaatgcaagcagctgacttcttcttctc 2244
 Db 2041 CTTCTCCCTCATGCTACCTACTGATCAATGCCAAAGCTGATGTCTTCTTATTTGCTTC 2100
 QY 2245 ccgagagcagagagctgtgctacatgagagagcagagcttcttgcagcagca 2304
 Db 2101 CCGGAGCGAGAGCTGTGATGATGAGAGAGCCCAAGCTTTTTCGACAGCTCAGGA 2160
 QY 2305 gttcaactctcagagcactgctcctcactgagacacacacaaatcttgcagtaact 2364
 Db 2161 GTTCACTTCCATGAGCATGCTGTCCACTGGGACACAACTCAATTTGTACTGGAACCT 2220
 QY 2365 ccttgagcctgctcagatgtggaatatacataaagagtggaactatccaagaagagag 2424
 Db 2221 CCTGGCGCTGCTCAGATTTGGATTAACATAAAGAGTGAAGTATCCAAAGAGACAG 2280
 QY 2425 gagaagaataaataatgctcctcctcctcctcctcctcctcctcctcctcctc 2484
 Db 2281 GGAGAAATTAATTAATGCTCCTCCTTCTCCTGCTTCCATCAGCCCTTCCCAT 2340
 QY 2485 ttcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2544
 Db 2341 TTTCTCTTCT 2400
 QY 2545 tttttaaactcctagttacccacacacacacacacacacacacacacacacacacac 2604
 Db 2401 TTTTAACTCCCTTACTTACCACTTATTTGCTTCAAGAGCCCTTCTTCTTCTTCT 2460
 QY 2605 cctgacctgactcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2664
 Db 2461 CCGGCCCTGATCT 2520
 QY 2665 tgggaagacataaagtttaagttaaggttcaaggttcaacatactgagccctgagagag 2724
 Db 2521 TGGGAAGACATTAAGGTTTATGTCAAGGTCAAACTTACATTAAGCCCTGAGAGAGG 2580
 QY 2725 ggcacatcctgagcctgagcctgctcctcctcctcctcctcctcctcctcctcct 2784
 Db 2581 GGCATCTCGGCTGAGCTACTGTCTCTTCCCACTGTCTTCTCAGGCTCTCAGAT 2640
 QY 2785 ggcacatagagtgagcgtgctgctgaggttggatatacaccacccagccagagtgctagt 2844
 Db 2641 GGCACATTAAGGGTGGCGGTGCGGGGGGGGATCCACCTCCAGCCACAGTCTCAGT 2700
 QY 2845 tfracttttataagcgttaatactatatttggtttggcttcttcttcttcttctt 2904
 Db 2701 TGTACTTTTATTAAGCTTATATCTATTTTGTGTTTGTCTTCTTCTTCTTCTTCTT 2760
 QY 2905 tglataatataatataatagatltcaatataatagattatccac 2964
 Db 2761 TGTAAATATATATATATATGATGATTAATTAATATATATATATATATATATATAT 2805

RESULT 3
 AF240469 2250 bp mRNA ROD 08-SEP-2000
 LOCUS AF240469 Mus musculus nicastrin mRNA, complete cds.
 DEFINITION AF240469
 ACCESSION AF240469
 VERSION AF240469.1 GI:9992879
 KEYWORDS house mouse.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2250)


```

Db 1318 ACATCTGCGGCGTGTCTGCGTACACCTGAGCTCTTCCACAATCGATTATACAGA 1377
1506 gatttaagacacactctgagacataaagtgatcattccgaatggcgaacccctgaag 1565
1378 GCATTTATGACACAGCGGTGAGAACTTAAATGACCTCTGAGTGCGAGAGCCAGAAAG 1437
1566 aggaactgaactctgttaacagacactcgaagccctgagatgtgagccaggtgctgg 1625
1438 AGGACCTCAACTTTGTGACAGACACCTGCCAAGGCAATGTGGCAGAGTCTGG 1497
1626 gaagtgctctgtatgagctgagagagaaacaaattcagcgacagctcagctgac 1685
1498 CGGTGACTGTATGAGCTGTGAGAGAAACCAACTTCAGACCTCCATCCAGAGCTGATC 1557
1686 cccaaggttaccgctctctatggtgtctgtatgaagcacaactatgtctcc 1745
1558 CCCAGACAGTTACAGCTCTCTATGAGTCTCTGTTAAAGCTAACACTACATCTGTTTC 1617
1746 agtctatcccaagcagagacaaagctcactgtggtgagcgggctctcacaattaca 1805
1618 AGTGATCTCGAAGAACATGACCTTAAGTCTATTGATGACAGAGCCTCTTCAACACTACA 1677
1806 tgcgtgtctcagagcccaacacacacactatgttgaagatgagctgagcaattga 1865
1678 TCGCCGTCTCCAGCCCTACACACAGACTTACGTTGAGTACGCTTGGCAACCTGA 1737
1866 ctggcagagtggttaacactcaccgagagagagtgagagagagagagagagagagag 1925
1738 CTGGCAGAGGAGGACCAACCTCAACCCGAGAGAGAGTCCAGATCCAAATGAATGCCAAATG 1797
1926 aaacaaagagatctgatagactcagtgagtgccagggcccttgcatctaaagagagc 1985
1798 AGAGCAAGAGATTTATTAATTAATCTGCGGAGTACAGGCCCTTGAAATTCACAGGAGAG 1857
1986 accgaactcccggtgtgtggtctcactgacagatlagcgaagggcctgtgtctctgct 2045
1858 AGAGGCTCCACAGTGTGTGCGTCCACAGTGCAGATGGCCAGGCGCTTGCCCTGCTCT 1917
2046 ttgaactgaagtcagtgagctctactgaatactctacatgagcagagagcggcgagag 2105
1918 TTGAATCTAGTACAGTGGAGCTCCACAGAAATCTTACCTGCGGAGAGAGCGGTGGAAG 1977
2106 atactcggcccgagatctctcactcagcagagagagagagagagagagagagagag 2165
1978 ACATCAAGAGCTGGATATCTTAATGTCACAGAAACCTTGAGTTCATCAGCGTGATCG 2037
2166 tgggtctgagacactcactctcctcactcagctgacactgaatgaatgaatgaatga 2225
2038 TGGGCTTCAGATCTTATCTCTCTCATCTGCTACCTGCTACCTGCTACCTGCTACCTG 2097
2226 atgtccttctatctctcccgagagagagagagagagagagagagagagagagagag 2285
2098 AGGTCTCTTTTGTCTCCCGAGAGCCAGAGCTGTGCTTACTGAAAGAGAGACTCTAGC 2157
2286 tttcttcagcagcagcagctcactcactagagatctgctcactcagcagcagcagc 2345
2158 TCTCCCTGCTCTCTGAACTTACTTCCACAGACAGGCTCGGCTGGAGAACAAACAC 2217
2346 taattgtcactggaactcctcctgagcctg 2375
2218 TAATTGTCACTGACTGTCTGTGGCCCTG 2247

```

```

RESULT 4
AL513282 155913 bp DNA HTG 26-FEB-2001
LOCUS Homo sapiens chromosome 1 clone RP11-179G5, *** SEQUENCING IN
DEFINITION
PROGRESS *** 8 unordered pieces.
ACCESSION AL513282.5 GI:13162044
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```

REFERENCE

AUTHORS

TITLE

JOURNAL

Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 155913)
Submitted (25-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquey@sanger.ac.uk
Requests: clone@sanger.ac.uk
On Feb 28, 2001 this sequence version replaced g1:13161719.
Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquey@sanger.ac.uk
Project Information
Center project name: ba179G5
Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 153938 bases at least 0.40
Consensus quality: 154365 bases at least 0.20
Insert size: 155213; sum-of-contrigs
Insert size: 164818; 1.0% error; agarose-fp
Quality coverage: 6.89x in Q20 bases; sum-of-contrigs quality
coverage: 7.27x in Q20 bases; agarose-fp

FEATURES

SOURCE

NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 20838: contrig of 20838 bp in length
* 20839 20938: gap of 100 bp
* 20939 50589: contrig of 29651 bp in length
* 50590 50689: gap of 100 bp
* 50690 81538: contrig of 30849 bp in length
* 81539 81638: gap of 100 bp
* 81639 89682: contrig of 8044 bp in length
* 89683 89782: gap of 100 bp
* 89783 108105: contrig of 18323 bp in length
* 108106 108205: gap of 100 bp
* 108206 149395: contrig of 41190 bp in length
* 149396 149495: gap of 100 bp
* 149496 153075: contrig of 3580 bp in length
* 153076 153175: gap of 100 bp
* 153176 155913: contrig of 2738 bp in length.
Location/Qualifiers
1. 155913
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-179G5"
/clone_id="RP11-179G5"
1. 20838
/note="assembly_fragment:02743
fragment_chain:1
clone_end:T7
vector_side:left"
20939. 50589
/note="assembly_fragment:01236
fragment_chain:1"
50690. 81538
/note="assembly_fragment:00478
fragment_chain:2"
81639. 89682
/note="assembly_fragment:03283
fragment_chain:2"
89783. 108105

Insert size: 167677; sum-of-contigs

Quality coverage: 131821; 35.28 error; agarose-fp

coverage: 4.70x in Q20 bases; sum-of-contigs quality

NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1
4822 4921: contig of 4821 bp in length
4922 11645: contig of 6728 bp in length
11646 11745: gap of 100 bp
11746 13929: contig of 2184 bp in length
13930 14029: gap of 100 bp
14030 16875: contig of 2846 bp in length
16876 16975: gap of 100 bp
16976 20388: contig of 3413 bp in length
20389 20488: gap of 100 bp
20489 29256: contig of 8768 bp in length
29257 29356: gap of 100 bp
30471 30570: contig of 1114 bp in length
30571 34082: gap of 100 bp
34083 34182: contig of 3512 bp in length
34183 49557: gap of 100 bp
49558 49657: contig of 15375 bp in length
49658 51777: contig of 2120 bp in length
51778 51877: gap of 100 bp
51878 53122: contig of 1245 bp in length
53123 53222: gap of 100 bp
53223 56791: contig of 3563 bp in length
56792 56891: gap of 100 bp
56892 61955: contig of 5064 bp in length
61956 62055: gap of 100 bp
62056 70347: contig of 8292 bp in length
70348 70447: gap of 100 bp
70448 77556: contig of 7109 bp in length
77557 77656: gap of 100 bp
77657 81110: contig of 3454 bp in length
81111 81210: gap of 100 bp
81211 82231: contig of 1021 bp in length
82232 82331: gap of 100 bp
82332 91668: contig of 9337 bp in length
91669 91768: gap of 100 bp
91769 112272: contig of 20504 bp in length
112273 112372: gap of 100 bp
112373 120545: contig of 8173 bp in length
120546 120645: gap of 100 bp
120646 140354: contig of 19709 bp in length
140355 140454: gap of 100 bp
140455 141779: contig of 1325 bp in length
141780 141879: gap of 100 bp
141880 143172: contig of 1293 bp in length
143173 143272: gap of 100 bp
143273 150964: contig of 7692 bp in length
150965 151064: gap of 100 bp
151065 152556: contig of 1492 bp in length
152557 152656: gap of 100 bp
152657 153707: contig of 1051 bp in length
153708 153807: gap of 100 bp
153808 154871: contig of 1064 bp in length
154872 154971: gap of 100 bp
154972 156380: contig of 1409 bp in length
156381 156480: gap of 100 bp
156481 160084: contig of 3604 bp in length
160085 160184: gap of 100 bp
160185 163849: contig of 3665 bp in length
163850 163949: gap of 100 bp

```

FEATURES * 163950 170677: contig of 6728 bp in length.

source

Location/Qualifiers

1. 170677

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="p11.1-13.3"

/clone="RP1-244G5"

/clone_1b="RPCT-1"

1. 4821

/note="assembly-fragment:01410"

fragment_chain:1"

4922. 11645

/note="assembly-fragment:00944"

fragment_chain:1"

11746. 13929

/note="assembly-fragment:00895"

fragment_chain:1"

14030. 16875

/note="assembly-fragment:01274"

fragment_chain:1"

16976. 20388

/note="assembly-fragment:00718"

fragment_chain:1"

20489. 29256

/note="assembly-fragment:00242"

fragment_chain:1"

29357. 30470

/note="assembly-fragment:00201"

fragment_chain:1"

30571. 34082

/note="assembly-fragment:01345"

fragment_chain:1"

34183. 49557

/note="assembly-fragment:01469"

fragment_chain:1"

49658. 51777

/note="assembly-fragment:01230"

fragment_chain:2"

51878. 53122

/note="assembly-fragment:00199"

fragment_chain:2"

53223. 56791

/note="assembly-fragment:01418"

fragment_chain:2"

56892. 61955

/note="assembly-fragment:01391"

fragment_chain:2"

62056. 70347

/note="assembly-fragment:01376"

fragment_chain:3"

70448. 77556

/note="assembly-fragment:00605"

fragment_chain:3"

77657. 81110

/note="assembly-fragment:00470"

fragment_chain:3"

81211. 82231

/note="assembly-fragment:00573"

fragment_chain:3"

82332. 91668

/note="assembly-fragment:01058"

fragment_chain:4"

91769. 112272

/note="assembly-fragment:00207"

fragment_chain:4"

112373. 120545

/note="assembly-fragment:01906"

fragment_chain:4"

120646. 140354

/note="assembly-fragment:00800"

fragment_chain:5"

Query Match	26.88;	Score 791;	DB 78;	Length 170677;
Best Local Similarity	99.48;	Pred. No. 2,4e+201;		
Matches 802; Conservative	2;	Mismatches 2;	Indels 1;	Gaps 1

Db	9713	TGCTGGGSGTGGGATCCACCTCCACCCACAGTCTCAGTTGACTTTTATTAAGCT	97172
Db	2863	gtaaatcctatttgcttggcttcttcctcttattccttttgtaataataataaat	2922
OY	9773	GTAATATCAATATTTTGTTTGCTTTTCTTCTTATCTTTTGAAAAATATATATAT	9832
Db	2923	gaattcattaaataagattatccac	2949
OY	9833	GAGTTTCATTAATAATGATTAATCCAC	9859

RESULT	7			PRI	29-APR-2000
LOCUS	AP001605				
DEFINITION	Homo sapiens genomic DNA, chromosome 21, clone:KB1987H1,				
ACCESSION	AP001605				
VERSION	AP001605.1	GI:7670559			
KEYWORDS	HTC.				
SOURCE	Homo sapiens pre-Pro-B cell cell_line:FLB14-14 DNA, clone_11b:Kelo9 RAC library clone:KB1987H1.				

ORGANISM	Homo sapiens	Chordata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa;		
Primates; Catarrhini; Hominae;		
Primates; Catarrhini; Hominae; Homo.		

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 165942)	Mammalia; Eutheria; Primates; Placentalia; Hominoidea; Hominidae; Homo	
Shimizu N., Kudoh J. and Shibuya K.		
Homo sapiens genomic DNA, chromosome 21, clone: KB1987H1,		

APP-D21S292 region
Published Only in Database (2000) In press
2 (bases 1 to 165942)

REFERENCE
AUTHORS Shimizu, N., Kudoh, J. and Shiduya, A.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2000) to the DDBJ/EMBL/GenBank databases.
network Shimizu, Keio University, School of Medicine, Molecular
100-8602 Japan

MOBAYUSHI SHINJUKU
Biologiy: 35 Shinnanomachi, Shinjuku-ku, Tokyo 160-0502, Japan
(E-mail: nshimizue@med.keio.ac.jp, Tel: 81-3-3351-2370,
Fax: 81-3-3351-2370)

```

FEATURES
source
Location/Qualifiers
1. .165942
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

/cell_line="FLEB14-14"
/cell_type="pre-pro-B cell"
/chromosome="21"
/clone="KB1987H1"

```

BASE COUNT	clone_11b="Kelo BAC library"
52737 a	30382 c 29628 g 53195 t

Query Match	11.38;	Score 332;	DB 91;	Length 165942;	!
Best Local Similarity	73.28;	Pred. No. 7.6e-78;			
		Mismatches 153;	Indels 71;	Gaps 1	

[illegible]

DG09786

Db 106343 TGAATTTGGCAGCTGTGATGCGTTCGGCTTTGGCAACCTCAACATTCTCCGAGAGGCAGGAGAGG
QY 2206 ctgcacatcagccaagactgatgtccctttcaattgctcccggcgagccaggagctgttc |
106343 106456

Db 106402 CTCCATCATGACAAACATGATGGTTTGTGCT-CCCTTAGCCAGGACGTGTTC 106403
3266 atactgaasgaaccacagctttctctgcagcctcagcagcttcaattcctctagagatctg 2325

2208
 106461
 2385

Qy 2326 TCCACAGGACATTAACCACTAGTTGTACACAGAACGTCCTCGACCTGCTCAGACTG 106558
 Db 106521 TCCACAGGACATTAACCACTAGTTGTACACAGAACGTCCTCGACCTGCTCAGACTG 106558

[illegible]

—

[illegible]


```

/misc_feature 1.4821
/obj xref="taxon:9606"
/chromosome="1"
/map "p11.1-13.3"
/clone "RPI-244G5"
/clone_lib="RPCT-1"
/clone="assembly-fragment:01410
fragment_chain:1"
/misc_feature 4922.11645
/clone="assembly-fragment:00944
fragment_chain:1"
/misc_feature 11746.13929
/clone="assembly-fragment:00895
fragment_chain:1"
/misc_feature 14030.16875
/clone="assembly-fragment:01274
fragment_chain:1"
/misc_feature 16976.20388
/clone="assembly-fragment:00718
fragment_chain:1"
/misc_feature 20489.29256
/clone="assembly-fragment:00242
fragment_chain:1"
/misc_feature 29357.30470
/clone="assembly-fragment:00201
fragment_chain:1"
/misc_feature 30571.34082
/clone="assembly-fragment:01345
fragment_chain:1"
/misc_feature 34183.49557
/clone="assembly-fragment:01469
fragment_chain:1"
/misc_feature 49658.51777
/clone="assembly-fragment:01230
fragment_chain:2"
/misc_feature 51878.53122
/clone="assembly-fragment:00199
fragment_chain:2"
/misc_feature 53223.56791
/clone="assembly-fragment:01418
fragment_chain:2"
/misc_feature 56892.61955
/clone="assembly-fragment:01391
fragment_chain:2"
/misc_feature 62056.70347
/clone="assembly-fragment:01376
fragment_chain:3"
/misc_feature 70448.77556
/clone="assembly-fragment:00605
fragment_chain:3"
/misc_feature 77657.81110
/clone="assembly-fragment:00470
fragment_chain:3"
/misc_feature 81211.82231
/clone="assembly-fragment:00573
fragment_chain:3"
/misc_feature 82332.91668
/clone="assembly-fragment:01058
fragment_chain:4"
/misc_feature 91769.112272
/clone="assembly-fragment:00207
fragment_chain:4"
/misc_feature 112373.120545
/clone="assembly-fragment:01906
fragment_chain:4"
/misc_feature 120646.140354
/clone="assembly-fragment:00800
fragment_chain:5"
/misc_feature 140455.141779
/clone="assembly-fragment:01052
fragment_chain:5"
/misc_feature 141880.143172
/clone="assembly-fragment:00033"

```

```

/misc_feature 143273.150964
/clone="assembly-fragment:00171"
/misc_feature 151065.152556
/clone="assembly-fragment:00257"
/misc_feature 152657.153707
/clone="assembly-fragment:01027"
/misc_feature 153808.154871
/clone="assembly-fragment:01448"
/misc_feature 154972.156380
/clone="assembly-fragment:01559"
/misc_feature 156481.160084
/clone="assembly-fragment:01800"
/misc_feature 160185.163849
/clone="assembly-fragment:00119
fragment_chain:6"
/misc_feature 163950.170677
/clone="assembly-fragment:01585
fragment_chain:6"

```

```

Query Match 7.2% Score 213.6; DB 78; Length 170677;
Best Local Similarity 98.2%; Pred. No. 5.5e-46;
Matches 216; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 10 ttcgcttgccctggaacacgaacttcggtctcttagctccggtccacagagagg 69
|||||
Db 166403 TTCGGAGAGCGCTCGAACAACAGACTTCGGTCTTAGGCTCCGGCCACAGAGACG 166344
QY 70 tgcagtgtagcctagagagccgcctaaacagacagagcggaggggttcgctga 129
|||||
Db 166343 TGTCCAGTGTAGGCTTAGAGGCGCGCTAACACAGAGAGCGGAACGGGGCTTCCTCA 166284
QY 130 gcaagagagcaagatgctcagcagcaggggtggtctcgtgagccgtagccggaaagtcggg 189
|||||
Db 166283 GCAGAGAGCGCAAGTGTGCTACGCGAGGGGTGCTGTGGGCTGACCCGGAGATCGGG 166224
QY 190 tctcctgcctctctgtcttcttcgtcgtcctcactagcaggt 229
|||||
Db 166223 TCTCCTTCGCTTCTGTCTTCTGCTGCTTACGAGGT 166184

```

```

RESULT 13
AL391417/c DNA HTG 24-JAN-2001
LOCUS Homo sapiens chromosome 6 clone RPI1-15J23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 45 unordered pieces.
ACCESSION AL391417
VERSION AL391417.4 GI:9944084
KEYWORDS HTG; HTGS_PHASEL.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 148540)
REFERENCE Sins,S.
AUTHORS Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 ISA, UK. E-mail enquiries: humquey@sanger.ac.uk
COMMENT requests: clonerequests@sanger.ac.uk
On Aug 29, 2000 this sequence version replaced gi:9931024.
Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquey@sanger.ac.uk
Project Information
Center project name: ba15j23
Summary Statistics
Sequencing program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 112851 bases at least Q40
Consensus quality: 127377 bases at least Q30
Consensus quality: 136503 bases at least Q20

```



```
misc_feature 1..6100
/note="assembly_fragment:02399
fragment_chain:1
clone_end:T7
vector_side:left"
6201..61918
/note="assembly_fragment:03300
fragment_chain:1"
62019..168537
/note="assembly_fragment:00055
fragment_chain:1
clone_end:SP6
vector_side:right"
misc_feature
BASE COUNT 53297 a 31342 c 32090 g 51608 t 200 others
ORIGIN
```

```

Query Match          6.1%; Score 180.8; DB 82; Length 168537;
Best Local Similarity 72.9%; Prod. No. 3./e-37;
Matches 366; Conservative 0; Mismatches 102; Indels 34; Gaps 9

QY 2461 ccttcccaatcacacctccttcgccattctccctctctctacacatcgagcatgttgagg 2520
DB 123236 CTTTTTCTGCTTTCCGTTCCTCCTCAATCGCTTAACCTCTCT--CCACCCCCTTGAGAACTGGGA 123179
QY 2521 tttaaatggaactctcttgctctgcgtttaactaccctagttaaccaacctaaattggccc 2580
DB 123178 TTACAAGAAGAAACACTCTCTGCTCCTCTCATAGTCCAGACACTATTAGTTACCCT 123119
QY 2581 tcaaga--ccctctacttttctctccgcgcctgtacactctctctgctccctaac-- 2634
DB 123118 TCATGATCCCTTCCTCACTCCTTCCTCAATCGCTCTATAAATCCTCTCTTCTCCACCCTCGS 123059
QY 2635 --ccacccccctgtaccacagacacactctcctaagtggagaag--gacataaaaggttaatatg 2690
DB 123058 CACCAACCCACTGTACCCAACCAACCTTCCTTACCAAGGAGAMAACAAGAGGGTTTAATGC 122999
QY 2691 cagaggtcaaac----tacattgagccccctgagga-----cagggagcactcttgagctg 2739
DB 122998 CAGAAATCAAACTATGTAACAATGAACTCTGAGAGAGAAACAGAGGGTGATTTCCTGGGCTG 122939
QY 2740 agacctactgtctctctccacacgtcctctctccagagccctcaagatgagcaattagagtg 2799
DB 122938 AGCCTGCTGTTTCTCTTGCCA--TTCTTTTCTCCAGGCTCCTCGAATGCCAATATCAATCAGTGS 122882
QY 2800 gcggtgctg--cgagtggtgatccacactccaagccacaagtgcacagtgtacttttat 2858
DB 122881 GTGTGCTGCGGGGTGGGCAATCCCACTCCAGGCCACAGTGTCAATTCACCTTTTAAATG 122822
QY 2859 agcgttaataatctatttgttttgtctctttctcttaattcttlttgtaataatata 2918
DB 122821 AGCTGTAATATCTATTTTTCCTGCTTCTTCTTCCATATTATTTT-----GTAAAAATG 122769
QY 2919 taatgagttcataataataga 2940
DB 122768 TACTGAGTTTATTPAAAATATA 122747

Search completed: June 27, 2001, 17:23:44
Job time: 10068 sec

```

Thu Jun 28 10:23:24 2001

us-09-541-094-13.rge


```

OY 2821 caccctcaagcccaagcagctcagctgagcttcttattatgaactgctgtaataatcattttgct 2880
    |||||||
Db 2821 caccctcaagcccaagcagctcagctgagcttcttattatgaactgctgtaataatcattttgct 2880
OY 2881 ttgtcttttcccttattctttcttgtaataataataataatgagttcatttaataatga 2940
    |||||||
Db 2881 ttgtcttttcccttattctttcttgtaataataataataatgagttcatttaataatga 2940
OY 2941 ttatccac 2949
    |||||||
Db 2941 ttatccac 2949

RESULT 2
AAA37886
ID AAA37886 standard; DNA: 2250 BP.
XX
XX AAA37886;
XX
XX 12-FEB-2001 (first entry)
XX
XX Mouse PAMP coding sequence.
XX
XX PAMP; mouse; presenilin associated membrane protein; immunogen;
XX neurodegenerative disease; Alzheimer's disease; lewy body variant;
XX parkinson's disease-dementia complex; neuropsychiatric disease;
XX schizophrenia; age-associated memory loss; developmental disorder;
XX neoplasm; diagnosis; ds.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 18..2144
XX /tag= a
XX /product= PAMP
XX
XX MO200060069-A1.
XX
XX 12-OCT-2000.
XX
XX 03-APR-2000; 2000MO-CA00354.
XX
XX 01-APR-1999; 99US-0127452.
XX 30-DEC-1999; 99US-0173826.
XX
XX (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX
XX St George-Hyslop PH, Fraser PE:
XX
XX MPI; 2000-665001/64.
XX P-PSDB; AAY97550.
XX
XX Isolated presenilin associated membrane proteins and nucleic acids
XX encoding them, useful for investigating and diagnosing Alzheimer's
XX disease and other neurodegenerative diseases -
XX
XX Claim 9; Page 70-71; 79pp; English.
XX
XX This sequence encodes the mouse presenilin associated membrane protein
XX (PAMP) of the invention. PAMP polypeptides may be used as an immunogen to
XX generate antibodies that recognise the PAMP polypeptide. The PAMP
XX nucleotide and protein sequence may also be used for diagnosing
XX individuals who are at risk or who have a variety of neurodegenerative
XX diseases (e.g. Alzheimer's disease, Lewy body variant, parkinson's
XX disease-dementia complex), neuropsychiatric disorders, and neoplasms.
XX Age-associated memory loss) to deduce the structural organisation and
XX These may further be used to deduce the structural organisation and
XX topology of PAMP, to identify proteins which interact with PAMP either
XX in concert with presenilin 1 (PS1) and PS2, or independently, and to
XX create cell-free systems, transfected cell lines, and animal models of
XX neurodegenerative and other diseases.
XX

```

QY	2821	caccctcagcccaagtgtctgaattgtaactttttaagcgtaataatcatatttgg	2880
Db	2821	caccttcagccccagtgctgcagttgtactttltaaaagcgttaataatccaatttgc	2880
QY	2881	ttagctttttccttatctcttttgtaaataataataatgaatgcattcatuaaataga	2940
Db	2881	tttgtcttttttccttattctcttttcgaaatatataataaaggattcattaaaataaga	2940
QY	2941	ttatcccac	2949
Db	2941	ttatcccac	2949

RESULT	2
AAA37886	
ID	AAA37886 standard; DNA; 2250 BP.

AC AAA37886;

DT 12-FEB-2001 (first entry)

Mouse PAMP coding sequence.

KW PAMP; mouse; presenilin ass

Parkinson's disease-dement.

KW neoplasm; diagnosis; ds-
 YX

XX	
OS	Mus sp

ET	EH	Key	CDS	Location/Qualifiers
			18.	2144

	E_{FT}	E_{IT}
/x lag= a		
/product= PAMP		

XX
PN
WO2000060069-A1

12-OCT-2000
PD
AA

03-APR-2000; 2000WO-CA00354

01-APR-1999; 99US-0127452

[illegible]

XX

[illegible]

DR P-PSDB; AAY97550

PT Isolated presenilin associated membrane proteins and nucleic acids

PT disease and other neurodegenerative diseases -

PS Claim 9; Page 70-71; 79pp; English.

CC This sequence encodes the mouse presentin associated membrane protein

CC generate antibodies that recognise the PAMP polypeptide. The PAMP

CC individuals who are at risk or who have a variety of neurodegenerative
CC diseases (e.g. Alzheimer's disease, Lewy body variant, Parkinson's

CC disease dementia complex), neuropsychiatric disorders, and neoplasms.

These may be used to deduce the structural organization and topology of PAMP. To identify proteins which interact with PAMP either

CC create cell-free systems, transfected cell lines, and animal models of

CC neurodegenerative and other diseases
XY

SQ Sequence 2250 BP; 535 A; 612 C; 563 G; 539 T; 1 other;

Query Match	59.2%;	Score 1746.4;	DB 21;	Length 2250;
Best Local Similarity	86.2%;	Pred. No. 0;		
Matches 1940;	Conservative 3;	Mismatches 304;	Indels 3;	Gaps 1.

QY	126	ctcagcagagagacgaagatgctcagcgacggggtgctgtctcgggctgcgacccgggaagt	185
Db	1	ccccgcggagagacgaacatgcacacgaacagggcgctcgggctgcgacccaggaattc	60
QY	186	ggggtctccttcgccttcgtctcttccttcgtcgtctactacgaggtttgtgcaggggaaact	245
Db	61	ggggctactctt---cttcgtccttttccgctgtgactgcagggatttggvbgggaaact	117
QY	246	cagtgvgagaggaagatatattatcccttaataaataaacagctccctgtgtcgcctgtcta	305
Db	118	cagtgvgagaggaataactcatctactcccttaataaataaacagctccctgtgtcgcctgtcta	177
QY	306	acgcacctatcatgatttggctcgcagcttcacattaaatgagagacaacggggttatccag	365
Db	178	acgcacctatcatgatttggctcgcagcttcacattaaatgagagatacaggggttatccag	237
QY	366	tagttagagaaagvgagggagccctacagctggtgattgactgtagtgcacaccccccttaaa	425
Db	238	tagttagagaaagaaagaaagacccgtaagctgggggtgtgtgtaacgtagtgcacaccccccttaaa	297
QY	426	tggctcgtcgtgagagcaagcattttacacagagatttaattgataaagctgaagggagaa	485
Db	298	tggctcgtcgtgagggcgaagctcttcacacagagatgtcaatgaggaagctgaagvgaca	357
QY	486	ccagcgcggaattcgtgtcttgcaagtgtctccttgacaagcccgagtcgtcgcagctct	545
Db	358	ccagtagaagatcgtgtcttgccgtgcagcttcacagacaagcccaactcaactcaagtctct	417
QY	546	ctctcctagtgtaagvgcccaaaatgtgaggtttgtgtttactccaattccattatgvgccag	605
Db	418	ctctcctagtgtagvgcccaaaatgtgaggtttgtgtattactccaactccctlaagvgccag	477
QY	606	agtttgctactgcagagaaatacacagtggaattcgtcgtggcacaatgtttgtgcttatgaag	665
Db	478	agtttgctactgcagagaaatacacatgtggaatgagtagtcggcacaagctgttgcttatgaag	537
QY	666	actttagtttcccacacttcttccttcgtgaagatgaataatgaaacccaagctcatcaagcagt	725
Db	538	actttagtttcccacacttcttccttcgtgaagatgagaaacccaagctcatcaagcagt	597
QY	726	gctatcaagatcacaccttgatgtagaattggtctcagacacaaccttccacattatgtgca	785
Db	598	gctatcaagatcacacaccttggatgcagaatgtgctctgcacaacaagcttcccatgtgtgtcta	657
QY	786	tgcagcctctttttacaacatgatatgtcttatcaagcactgcacacctgcagatgcggcagct	845
Db	658	tgcagcctctcttcacaacatgatacgcgttatcaagcactgcacacctgcagatgcggcagct	717
QY	846	ccatccaaagacacttcaagacaacaacccaanaatgtctcgttacccccgttcgttatca	905
Db	718	tcatccagagacacttcaagacaacaacccaanaatgtctcgttacccccgttcgttatca	777
QY	906	atgtgtgtagacatgtcttaagcctataataacaacatgvgagacattaaagccttgacagacag	965
Db	778	acgtatgtgagacatgtcttaagcctataataacatctgtgagactagaactgaactcagtcaggg	837
QY	966	tgtgtgtgtgcgcacccgcgttgatgtagtgcgttcccttttcggagatgtggccccaggg	1025
Db	838	tgtgtgtgtgcgcgcacccgcgttgatgtagcgcgttcccttttcggagatgtggccccaggg	897
QY	1026	ctgaagagcgcaagtgtctccttgttcaaccacagctgcgtcgtcgtgaacatttgacaag	1085
Db	898	ctgaagatgtgtcgttagcctcctcttgcataccagctgcgtcgaagcttttgacaag	957
QY	1086	caactgtgtgacacacccctggcccccgaaatgtcaatgttgccttctcttcaaggggaact	1145

Db	958	caccgcatgtgcacactctatccacgaaatgtgaatgtttcttcttccaaaggggaaactt	1017
QY	1146	ttgacttaacttggcagctcgagatagtgctacgaatgtgagaagggcaagttccggc	1205
Db	1018	ttgacttaacttggcagctcgagatagtgctacgaatgtgagaagggcaagttccggc	1077
QY	1206	agttgagaatgttgaactcaatttgtgagcttggagcaggtggccttaagaacttcatag	1265
Db	1078	ggctcgagaagaactgcacactctcttctgtgagctgtggagacaggtggcccttaagaacttcatag	1137
QY	1266	agcttctgatatgcacacaaatccctgtttcttcaagaaagatctgtacgcgaacaggttgg	1325
Db	1138	atctctgatatgcacacaaatccctgtttcttcaagaaagatctgtacgcgaacaggttgg	1197
QY	1326	aggaatccctggccaacttggagaagaatgtgtgtgtgtccctgtctgtcaactccagga	1385
Db	1198	aggaatctctgtggccaacttggagaagaatgtgtgtgtgtccctgtctgtcaactccagga	1257
QY	1386	ggccaactcattcccaagctcttcccaacttctccctggcagagatcttcttcagactcga	1445
Db	1258	gacgtggccaactcctcccaagctcttcccaacttctcccaagagatcttcttcagactcga	1317
QY	1446	acaatctctggagttgttctgtgtcgaacacttgggtgtgtcttccatcaaaaataatccaga	1505
Db	1318	acaatctctggagttgttctgtgtcgaacacttgggtgtgtcttccatcaaaaataatccaga	1377
QY	1506	gtatttacgcacactgtctgagagaacataatgttgagctatccccaatgtgtcgaacctga	1565
Db	1378	gcattttatgacacgctctgagagaacataatgttgagctatccccaatgtgtcgaacctga	1437
QY	1566	aggacactgaaattgttaacagacacacttgcgaagccctggcagatgtgtgccaggtgtcgg	1625
Db	1438	aggacactcgaacttgtgtacagaaacttgcgaagccctggcagatgtgtgccaggtgtcgg	1497
QY	1626	gacgtgtctgtatgagcttgcagaggaagaaaccaacttcaacgcgacaaagtccaagctgac	1685
Db	1498	cgcgtgacgtgtatgagcttgcagaggaagaaaccaacttcaacgcgacaaagtccaagctgac	1557
QY	1686	cccaacaggttaccgcgcgtctctatgtgttctctgtatlaaagccaacactcagtgttcc	1745
Db	1558	cccaagacagttaacagcttctctatgtgttctctgtatlaaagccaacactcagtgttcc	1617
QY	1746	agctatccctcagggcagagacttaagttcctacttgggtgagcggccttccaacttaca	1805
Db	1618	agctcagatccctcagaaacatlgacacttaagttcctacttgggtgagcggccttccaacttaca	1677
QY	1806	tgcgtgtctcccaagccccaacaaacacactatgttgaagatagtgcttgggaaatttga	1865
Db	1678	tgcgcgtctcccaagcccttaccacaagacacttaagttgtgtgacatgacgcttgggaaacctga	1737
QY	1866	cttggcacaagtgtgtcaacctcaccggaacggagagatgtgccagagatccaagtlaaagtc	1925
Db	1738	cttggcacaaggtgacccaacctcaccggaacggagagatgtgccagagatccaagtlaaagtc	1797
QY	1926	aaaaaagagatctgtatgtgtactatcatgtgttccagggcccttgcattcctaagagacgg	1985
Db	1798	agagcagaagatataatgaataactctgtgtgttacaagggcccttgcattcctaagagacgg	1857
QY	1986	accggaatcccccggtgtgtgtgcgttttcaactgcagacagattgacgggccttgtctctgcct	2045
Db	1858	agagaggtccccaacaggtgtgtgtgcgttccacaaggtgtgcagggccttgtctctgcct	1917
QY	2046	ttgaactgtacgtatgtgagcttactgaatactacatactgagctatagagccgctgtgaag	2105
Db	1918	ttgaactgtacgtatgtgagcttccacagataactctacgttggcggagagcgctgtgaag	1977
QY	2106	atataccgtggccgggatatattctatcgcgcgaagaaagacttgaatgtgacccctgaag	2165
Db	1978	aatcccaagctctggaatacttccatattgtccagaanaaagcttgaatgtgacccctgaag	2037
QY	2166	tgggttctgggaactcgaatcttcttccctcattcgttactactatgtacaatgtccaagctgtg	2225
Db	2038	tgggttctgaagatcccttatcttctctcccaatcgttcaactactatgtacaatgtccaagctgtg	2097

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 1: SEQ ID 140; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

SO Sequence 424 BP; 98 A; 104 C; 121 G; 101 T; 0 other;

Query Match 14.4%; Score 424; DB 21; Length 424;
 Best Local Similarity 100.0%; Pred. No. 1.2e-107;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 agagaggaagaatgctacgcaagggtgtgctgtgagctgacccgggaatcggtgctc 191
 Db 1 agagaggaagaatgctacgcaagggtgtgctgtgagctgacccgggaatcggtgctc 60
 QY 192 tccctgcctctgtcttctgcgtccctacagcaggtgtgcaagggaatcagctgg 251
 Db 61 tccctgcctctgtcttctgcgtccctacagcaggtgtgcaagggaatcagctgg 120
 QY 252 agaggaatataatatacccttaataaacaagctccctgtgtgctgctcaacgcca 311
 Db 121 agaggaatataatatacccttaataaacaagctccctgtgtgctgctcaacgcca 180
 QY 312 ctcatcaatgctgctgacgtcttcaatagtagagacagaggttatccagctagtag 371
 Db 181 ctcatcaatgctgctgacgtcttcaatagtagagacagaggttatccagctagtag 240
 QY 372 agaaagagagacctaagctggtgattgactgagtgcccaaccccttaacgtgttc 431
 Db 241 agaaagagagacctaagctggtgattgactgagtgcccaaccccttaacgtgttc 300
 QY 432 tgcctgagagagacatttaccagagattatgaggaagctggaagaggaacagcc 491
 Db 301 tgcctgagagagacatttaccagagattatgaggaagctggaagaggaacagcc 360
 QY 492 gaattgctgtcttgcagctgcttgcacaaagccagctcctgctcagagcttctccta 551
 Db 361 gaattgctgtcttgcagctgcttgcacaaagccagctcctgctcagagcttctccta 420
 QY 552 gtgt 555
 Db 421 gtgt 424

RESULT 5

AAA37890
 ID AAA37890 standard; DNA; 473 BP.

AC AAA37890;

XX 12-FEB-2001 (first entry)

XX EST sequence #2.

XX PAMP; presenilin associated membrane protein; immunogen;

KW neurodegenerative disease; Alzheimer's disease; Lewy body variant;
 KW Parkinson's disease-dementia complex; neuropsychiatric disease;
 KW schizophrenia; age-associated memory loss; developmental disorder;
 KW neoplasm; diagnosis; EST; expressed sequence tag; ds.
 OS Unidentified.

XX W0200060069-A1.

XX 12-OCT-2000.

XX 03-APR-2000; 2000WO-CA00354.

XX 01-APR-1999; 99US-0127452.

XX 30-DEC-1999; 99US-0173826.

XX (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX St George-Hyslop PH, Fraser PE;

XX WPI; 2000-665001/54.

XX Isolated presenilin associated membrane proteins and nucleic acids
 PT encoding them, useful for investigating and diagnosing Alzheimer's
 PT disease and other neurodegenerative diseases -

XX Example 1: Page 60; 79pp; English.

CC This sequence represents an expressed sequence tag used to isolate the
 CC human presenilin associated membrane protein (PAMP) coding sequences
 CC of the invention. PAMP polypeptides may be used as an immunogen to
 CC generate antibodies that recognise the PAMP polypeptide. The PAMP
 CC nucleotide and protein sequence may also be used for diagnosing
 CC individuals who are at risk or who have a variety of neurodegenerative
 CC diseases (e.g. Alzheimer's disease, Lewy body variant, Parkinson's
 CC age-associated memory loss), neuropsychiatric disorders (e.g. schizophrenia,
 CC These may further be used to deduce the structural organisation and
 CC topology of PAMP, to identify proteins which interact with PAMP either
 CC in concert with presenilin 1 (PS1) and PS2, or independently, and to
 CC create cell-free systems, transfected cell lines, and animal models of
 CC neurodegenerative and other diseases.

SO Sequence 473 BP; 129 A; 116 C; 117 G; 111 T; 0 other;

Query Match 12.1%; Score 357.8; DB 21; Length 473;
 Best Local Similarity 84.8%; Pred. No. 2.9e-89;
 Matches 401; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 249 tggagaggaagatataatcccttaataaacaagctccctgtgtgctgctcagc 308
 Db 1 tggagaggaagatataatcccttaataaacaagctccctgtgtgctgctcagc 60
 QY 309 ccactcatagattgtgctgctcagcttcaatagtagagacaggggtatccagtag 368
 Db 61 ccactcatagattgtgctgctcagcttcaatagtagagacaggggtatccagtag 120
 QY 369 tagagaaagagagacctaagctggtgattgactgagtgcccaaccccttaacgtg 428
 Db 121 tagagaaagagagacctaagctggtgattgactgagtgcccaaccccttaacgtg 180
 QY 429 ttctctgagagagacatttaccagagattatgaggaagctggaagaggaaccca 488
 Db 181 ttctctgagagagacatttaccagagattatgaggaagctggaagaggaaccca 240
 QY 489 gccgaattgctggtcttgcagctgcttgcacaaagccagctcgtcaggtctctc 548
 Db 241 gtagaattgctggtcttgcagctgcttgcacaaagccagctcgtcaggtctctc 300
 QY 549 ctatgtacagtgcccaatgatgtgtgtgttactccaattccttagggccagag 608
 Db 301 ctatgtacagtgcccaatgatgtgtgtgttactccaattccttagggccagag 360

QY 609 ttgtcactgcagagaatacagtggaatcgcgtggcaatgtgttgcttatgaagact 668
|||||
Db 361 ttgtcactgcagagaatacagtggaatcgcgtggcaatgtgttgcttatgaagacc 420

QY 669 ttgttcccccactcttctcttggaatgaataatgaaccgaagcacaag 721
|||||
Db 421 ttgttcccccactcttctcttggaatgaataatgaaccgaagcacaag 473

RESULT 6

AAA37891
ID AAA37891 standard; DNA; 463 BP.

AAA37891;

12-FEB-2001 (first entry)

EST sequence #3.

PAMP; presenilin associated membrane protein; immunogen;
neurodegenerative disease; Alzheimer's disease; Lewy body variant;
Parkinson's disease-dementia complex; neuropsychiatric disease;
schizophrenia; age-associated memory loss; developmental disorder;
neoplasm; diagnosis; EST; expressed sequence tag; ds.

Unidentified.

WO200060069-A1.

12-OCT-2000.

03-APR-2000; 2000WO-CA00354.

01-APR-1999; 99US-0127452.

30-DEC-1999; 99US-0173826.

(UTOR) UNIV TORONTO GOVERNING COUNCIL.

St George-Hyslop PH, Fraser PE;

WPI; 2000-665001/64.

Isolated presenilin associated membrane proteins and nucleic acids
encoding them, useful for investigating and diagnosing Alzheimer's
disease and other neurodegenerative diseases -

Example 1; Page 60; 79pp; English.

This sequence represents an expressed sequence tag used to isolate the
human presenilin associated membrane protein (PAMP) coding sequences
of the invention. PAMP polypeptides may be used as an immunogen to
generate antibodies that recognize the PAMP polypeptide. The PAMP
nucleotide and protein sequence may also be used for diagnosing
individuals who are at risk or who have a variety of neurodegenerative
diseases (e.g. Alzheimer's disease, Lewy body variant, Parkinson's
disease-dementia complex), neuropsychiatric diseases (e.g. schizophrenia,
age-associated memory loss), developmental disorders, and neoplasms.
These may further be used to deduce the structural organisation and
topology of PAMP, to identify proteins which interact with PAMP either
in concert with presenilin 1 (PS1) and PS2, or independently, and to
create cell-free systems, transfected cell lines, and animal models of
neurodegenerative and other diseases.

Sequence 463 BP; 113 A; 134 C; 113 G; 103 T; 0 other;

Query Match

Best Local Similarity 86.2%; Pred. No. 3,7e-87;

Matches 399; Conservative 0; Mismatches 63; Indels 1; Gaps 1;

QY 1437 gagctcgaacatctctggtgttctgtgacccactgtgtccttcataacaat 1496
|||||

Db 1 gggctcgaacatctctggtgttctgtgacccactgtgtccttcataacatcgt 60
QY 1497 attaccagagattattagacactctgtgagaacattatgtatgctatccgaatgctga 1556
|||||
Db 61 attaccagagattattagacactctgtgagaacattatgtatgctatccgaatgctga 120
QY 1557 gccctcgaagagaccttgactttgtaacagacactgtgacccctgtgagatgtggca 1616
|||||
Db 121 gccatgaagaagaccctcaacttctgtgacagacactgtgacccctgtgagatgtggca 180
QY 1617 cgtgtcgtgagacgtctctgtatgactgtcagagaggaacacttcagacagatgc 1676
|||||
Db 181 cagtgtcgtgagacgtctctgtatgactgtcagagaggaacacttcagacagatgc 240
QY 1677 aggtctgaccccaacagtgatccgcctgtctatgggttccgtatgaacccaacact 1736
|||||
Db 241 aggtctgaccccaacagtgatccgcctgtctatgggttccgtatgaacccaacact 300
QY 1737 catgttccagctctatccagcagacactaagttcctactgtgtgagagcctctc 1796
|||||
Db 301 catgttccagctctatccagcagacactaagttcctactgtgtgagagcctctc 360
QY 1797 aacattacatcgtctctccagcccaacacacactatgtgtacagatgctcttg 1856
|||||
Db 361 aacattacatcgtctctccagcccaacacacactatgtgtacagatgctcttg 420
QY 1857 caatttgcact-ggacacagtggtcaactcaccgagagcagt 1898
|||||
Db 421 caaactgactgtgcaagggacacacactcaccgagagcagt 463

RESULT 7

AAA37892
ID AAA37892 standard; DNA; 481 BP.

AAA37892;

12-FEB-2001 (first entry)

EST sequence #4.

PAMP; presenilin associated membrane protein; immunogen;
neurodegenerative disease; Alzheimer's disease; Lewy body variant;
Parkinson's disease-dementia complex; neuropsychiatric disease;
schizophrenia; age-associated memory loss; developmental disorder;
neoplasm; diagnosis; EST; expressed sequence tag; ds.

Unidentified.

WO200060069-A1.

12-OCT-2000.

03-APR-2000; 2000WO-CA00354.

01-APR-1999; 99US-0127452.

30-DEC-1999; 99US-0173826.

(UTOR) UNIV TORONTO GOVERNING COUNCIL.

St George-Hyslop PH, Fraser PE;

WPI; 2000-665001/64.

Isolated presenilin associated membrane proteins and nucleic acids
encoding them, useful for investigating and diagnosing Alzheimer's
disease and other neurodegenerative diseases -

Example 1; Page 61; 79pp; English.

This sequence represents an expressed sequence tag used to isolate the
human presenilin associated membrane protein (PAMP) coding sequences
of the invention. PAMP polypeptides may be used as an immunogen to

QY 306 accgcatcatcagattggtcctcagcttcaattagttgagacacaggggttaccag 365
 |||||
 Db 178 agcccatcatcagattggtcctcagcttcaattagttgagacacaggggttaccag 237
 QY 366 tagtgaagaagaagagagacagctggttactgactgagcccccacttaca 425
 |||||
 Db 238 tagtgaagaagaagaaga-ctgaagtggtgttgacagtggtcccaaccccttactg 296
 QY 426 tggctcgtcgtgagagcaagcatttaccagagattatgagagaagctga 476
 |||||
 Db 297 gctcgtcgtgagagag-----cttcacagagatgtaattgagagaagctga 341

RESULT 11

AAV86828
 ID AAV86828 standard; cDNA: 247 BP.

AC AAV86828;

DT 27-APR-1999 (first entry)

DE EST clone BB54.

Expressed sequence tag; secreted protein; haematopoiesis regulator;
 tissue growth; actinin; inhibin; tumour invasion suppressor; EST; human;
 chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

OS Homo sapiens.

PN MO9845435-A2.

PD 15-OCT-1998.

PF 10-APR-1998: 98WO-US06954.

PR 10-APR-1997: 97US-0835913.

PA (GEMV) GENETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racine LA, Spaulding V, Treacy M;

DR WPI: 1999-070076/06.

PT New polynucleotides encoding human secreted proteins - derived from
 e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries

PS Claim 1; Page 361; 633pp; English.

CC This sequence represents an expressed sequence tag (EST), and is a
 CC polynucleotide of the invention. The polynucleotides of the invention are
 CC all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.

SO Sequence 247 BP; 62 A; 60 C; 53 G; 72 T; 0 other;

Query Match 6.3%; Score 187; DB 20; Length 247;

Best Local Similarity 100.0%; Pred. NO. 5e-42;

Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 578 ggtttactccaattccatgagggccagagtttgcctcactgcagagaatacagtggaat 637
 |||||
 Db 56 ggtttactccaattccatgagggccagagtttgcctcactgcagagaatacagtggaat 115
 QY 638 tgcgtggccaatggttggcttatgaagacttaagttcccatcttcttcttgaagat 697
 |||||
 Db 116 tgcgtggccaatggttggcttatgaagacttaagttcccatcttcttcttgaagat 175
 QY 698 gaaatgaacccaagatcatcagcagtgctatcaagatcacacactgagtaagtc 757
 |||||
 Db 176 gaaatgaacccaagatcatcagcagtgctatcaagatcacacactgagtaagtc 235
 QY 758 tcagcac 764
 |||||
 Db 236 tcagcac 242

RESULT 12

AAA37894
 ID AAA37894 standard; DNA: 172 BP.

AC AAA37894;

DT 12-FEB-2001 (first entry)

DE EST sequence #6.

PAMP; presenilin associated membrane protein; immunogen;
 neurodegenerative disease; Alzheimer's disease; Lewy body variant;
 Parkinson's disease-dementia complex; neuropsychiatric disease;
 schizophrenia; age-associated memory loss; developmental disorder;
 neoplasm; diagnosis; EST; expressed sequence tag; ds.

OS Unidentified.

PN WO200060069-A1.

PD 12-OCT-2000.

PF 03-APR-2000: 2000WO-CA00354.

PR 01-APR-1999: 99US-0127452.

PR 30-DEC-1999: 99US-0173826.

PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.

PI St George-Hyslop PH, Fraser PE;

DR WPI: 2000-665001/64.

PT Isolated presenilin associated membrane proteins and nucleic acids
 PT encoding them, useful for investigating and diagnosing Alzheimer's
 PT disease and other neurodegenerative diseases -

PS Example 1; Page 61-62; 79pp; English.

CC This sequence represents an expressed sequence tag used to isolate the
 CC human presenilin associated membrane protein (PAMP) coding sequences
 CC of the invention. PAMP polypeptides may be used as an immunogen to
 CC generate antibodies that recognise the PAMP polypeptide. The PAMP
 CC nucleotide and protein sequence may also be used for diagnosing
 CC individuals who are at risk or who have a variety of neurodegenerative
 CC diseases (e.g. Alzheimer's disease, Lewy body variant, Parkinson's
 CC disease-dementia complex), neuropsychiatric diseases (e.g. schizophrenia,
 CC age-associated memory loss), developmental disorders, and neoplasms.
 CC These may further be used to deduce the structural organisation and
 CC topology of PAMP, to identify proteins which interact with PAMP either
 CC in concert with presenilin 1 (PS1) and PS2, or independently, and to
 CC create cell-free systems, transfected cell lines, and animal models of
 CC neurodegenerative and other diseases.

SO Sequence 172 BP; 37 A; 46 C; 50 G; 39 T; 0 other;

Query Match	4.1%;	Score 120.6;	DB 21;	Length 425;
Best Local Similarity 65.3%;	Pred. No. 1.0e-23;			
Matches 241;	Conservative 0;	Mismatches 119;	Indels 9;	Gaps 4.
QY 2578	ccttagagaccccttctacttcttctctctgcccgtacccctctctctgctccctaccccc	2637		
DB 395	CCCTTGTGCTTCATCTCTGCCCCCTTACTGCCCTCTGTGACTCTGACCTATGTACCCGACACCC	336		
QY 2638	accctcgtaaccagacccttctctgtagga--agacataaaaggttaatgtaagg	2695		
DB 335	TCCCCCATATTCAGCCACCTGCTGCACGACGAAAGGGGTGTGAAGAATTGTGATCTGCA	276		
QY 2696	tcaactacattgaagcccttgaggagagagatctctggtctgagcctactgtctctt	2755		
DB 275	TTCAACTACCTTAACTTAAGGAAGAATGGATTCCCTGGCTGACGACACTGTCTCTT	216		
QY 2756	cccactgtccttcttccagagccctcagatgacataaggttgagtcgtctgtc-gagtg	2814		
DB 215	CCCACTGTCTCTTCTCCAGAGTGTCAGATGGCATGTGTATGTGGGACAGCTGTAACTGG	156		
QY 2871	tattttgttttctgtctt--cccttattcttttggtaatatataataatgagttc	2928		
DB 95	TATTTTGTGGGGTTTTTTCCTTTTCTTTTGTGTAATATATATATATATATATATAT	36		
QY 2929	cattnaat 2937			
DB 35	CATTAAAT 27			
RESULT 14				
AAA37887				
ID AAA37887	standard; DNA; 2942 BP.			
AC AAA37887;				
DT 12-FEB-2001	(first entry)			
DE D. melanogaster PAMP coding sequence.				
XX PAMP; presenilin associated membrane protein; immunogen;				
KW neurodegenerative disease; Alzheimer's disease; Lewy body variant;				
KW Parkinson's disease-dementia complex; neuropsychiatric disease;				
KW schizophrenia; age-associated memory loss; developmental disorder;				
XX neoplasm; diagnosis; ds.				
OS Drosophila melanogaster.				
XX Key	Location/Qualifiers			
EH CDS	83..2170			
FT	/*tag= a			
FT	/product= PAMP			
XX MO20006069-A1.				
XX PD 12-OCT-2000.				
XX PF 03-APR-2000; 2000WO-CA00354.				
XX PR 01-APR-1999; 99US-0127452.				
XX PR 30-DEC-1999; 99US-0173826.				
XX PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.				
XX PI St George-Hyslop PH, Fraser PE;				
XX MPI: 2000-665001/64.				
DR P-PSDB; AAY97551.				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 27, 2001, 14:51:51 ; Search time 86.97 Seconds
(without alignments)
6281.853 Million cell updates/sec

Title: US-09-541-094-13

Perfect score: 2949

Sequence: 1 tctgagatcgcgttcg.....attaatagatattccac 2949

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/1na/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/1na/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	79.2	2.7	7218 1	US-08-232-463-14 Sequence 14, Appl
2	49.8	1.7	53526 3	US-081658-136-2 Sequence 2, Appl
3	49.8	1.7	53577 3	US-081658-136-1 Sequence 1, Appl
4	45.4	1.5	4698 1	US-07-807-043B-5 Sequence 5, Appl
5	45.4	1.5	4698 2	US-08-299-849B-5 Sequence 5, Appl
6	45.4	1.5	4698 3	US-08-142-368A-5 Sequence 5, Appl
7	45.4	1.5	4698 4	US-08-967-727-5 Sequence 5, Appl
8	45.4	1.5	4698 5	US-08-037-230D-5 Sequence 5, Appl
9	41.4	1.4	289 1	US-08-341-568-3 Sequence 3, Appl
10	41.4	1.4	289 2	US-08-911-020-3 Sequence 3, Appl
11	41.4	1.4	5852 1	US-07-867-106-2 Sequence 2, Appl
12	40.6	1.4	2409 4	US-09-293-322C-8 Sequence 8, Appl
13	40.2	1.4	19124 4	US-08-487-826B-13 Sequence 13, Appl
14	39.8	1.3	1117 4	US-09-247-373B-33 Sequence 33, Appl
15	39.6	1.3	1736 4	US-09-182-816-22 Sequence 22, Appl
16	39.6	1.3	1736 5	US-09-182-816-24 Sequence 24, Appl
17	39.6	1.3	1736 6	US-09-471-528-22 Sequence 22, Appl
18	39.6	1.3	1736 7	US-09-471-528-24 Sequence 24, Appl
19	39.4	1.3	11517 4	US-07-920-281C-1 Sequence 1, Appl
20	39.4	1.3	11517 5	US-08-466-277-1 Sequence 1, Appl
21	39.2	1.3	1582 3	US-08-545-196B-10 Sequence 10, Appl
22	39.2	1.3	1582 4	US-08-545-196B-12 Sequence 12, Appl
23	39.2	1.3	1582 5	US-09-112-096-14 Sequence 14, Appl
24	39.2	1.3	1582 6	US-09-112-096-16 Sequence 16, Appl
25	38.8	1.3	1641 1	US-08-300-903A-8 Sequence 8, Appl
26	38.4	1.3	960 3	US-09-248-335-57 Sequence 57, Appl
27	38.2	1.3	2852 3	US-09-027-137-2 Sequence 2, Appl

C 28	38	1.3	688	6	5498694-3	Patent No. 5498694
C 29	38	1.3	1493	1	US-08-340-820-24	Sequence 24, Appl
C 30	38	1.3	1493	1	US-08-593-535-24	Sequence 24, Appl
C 31	37.6	1.3	140	1	US-08-628-417-5	Sequence 5, Appl
C 32	37.6	1.3	240	1	US-08-628-417-6	Sequence 6, Appl
C 33	37.6	1.3	2294	4	US-08-964-700A-1	Sequence 1, Appl
C 34	37.6	1.3	7938	4	US-09-331-581-3	Sequence 3, Appl
C 35	37.6	1.3	7938	4	US-09-331-581-14	Sequence 14, Appl
C 36	37.4	1.3	445	2	US-08-332-766A-16	Sequence 16, Appl
C 37	37.4	1.3	746	4	US-09-013-810-1	Sequence 1, Appl
C 38	37.4	1.3	2230	4	US-08-378-313-24	Sequence 24, Appl
C 39	37.4	1.3	3238	5	PCT-US94-10080-5	Sequence 5, Appl
C 40	37.4	1.3	5173	1	US-08-242-677-1	Sequence 1, Appl
C 41	37.4	1.3	5852	1	US-07-867-106-2	Sequence 2, Appl
C 42	37.2	1.3	966	2	US-08-766-738-2	Sequence 2, Appl
C 43	37	1.3	1474	4	US-08-821-994-64	Sequence 64, Appl
C 44	36.8	1.2	1441	4	US-08-821-994-65	Sequence 65, Appl
C 45	36.4	1.2	271	2	US-08-731-272A-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOMLOPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)663-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT9gpt-Fls
US-08-232-463-14

RESULT 8
 US-08-037-230D-5
 Sequence 5, Application US/08037230D
 Patent No. 6235525
 GENERAL INFORMATION:
 APPLICANT: Gaugler, B attrice; Van den Eynde, Beno t;
 Applicant: van der Bruggen, Pierre; Boon-Fallieur, Thierry
 TITLE OF INVENTION: Isolated Nucleic Acid Molecules Coding For
 TITLE OF INVENTION: Tumor Rejection Antigen Precursor Mage-3 And Uses Thereof
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESS: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 COMPUTER: IBM
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/037,230D
 FILING DATE: 26-MARCH-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/04354
 FILING DATE: 22-MAY-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/807,043

[illegible]

RESULT 9
 US-08-341-568-3/C
 Sequence 3, Application US/08341568
 Patent No. 5661021
 GENERAL INFORMATION:
 APPLICANT: Buchert, Johanna
 APPLICANT: Silka-aho, Matti
 APPLICANT: Viikari, Liisa
 APPLICANT: Penttila, Merja
 APPLICANT: Saloheimo, Anu
 APPLICANT: Marjatta, Rannu
 TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
 TITLE OF INVENTION: methods for isolating the genes, and methods
 TITLE OF INVENTION: pulps
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Birch, Stewart, Kolasch and Birch
 STREET: PO Box 747
 CITY: Falls Church
 STATE: VA
 COUNTRY: US
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible

[illegible]

Query Match	1.38;	Score 39.8;	DB 4;	Length 1117;
Best Local Similarity	62.68;	Pred. No. 0.19;		
Matches 62; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0

RESULT 15
US-09-182-816-22/c
; Sequence 22, Application US/09182816
; Patent No. 6143542

```

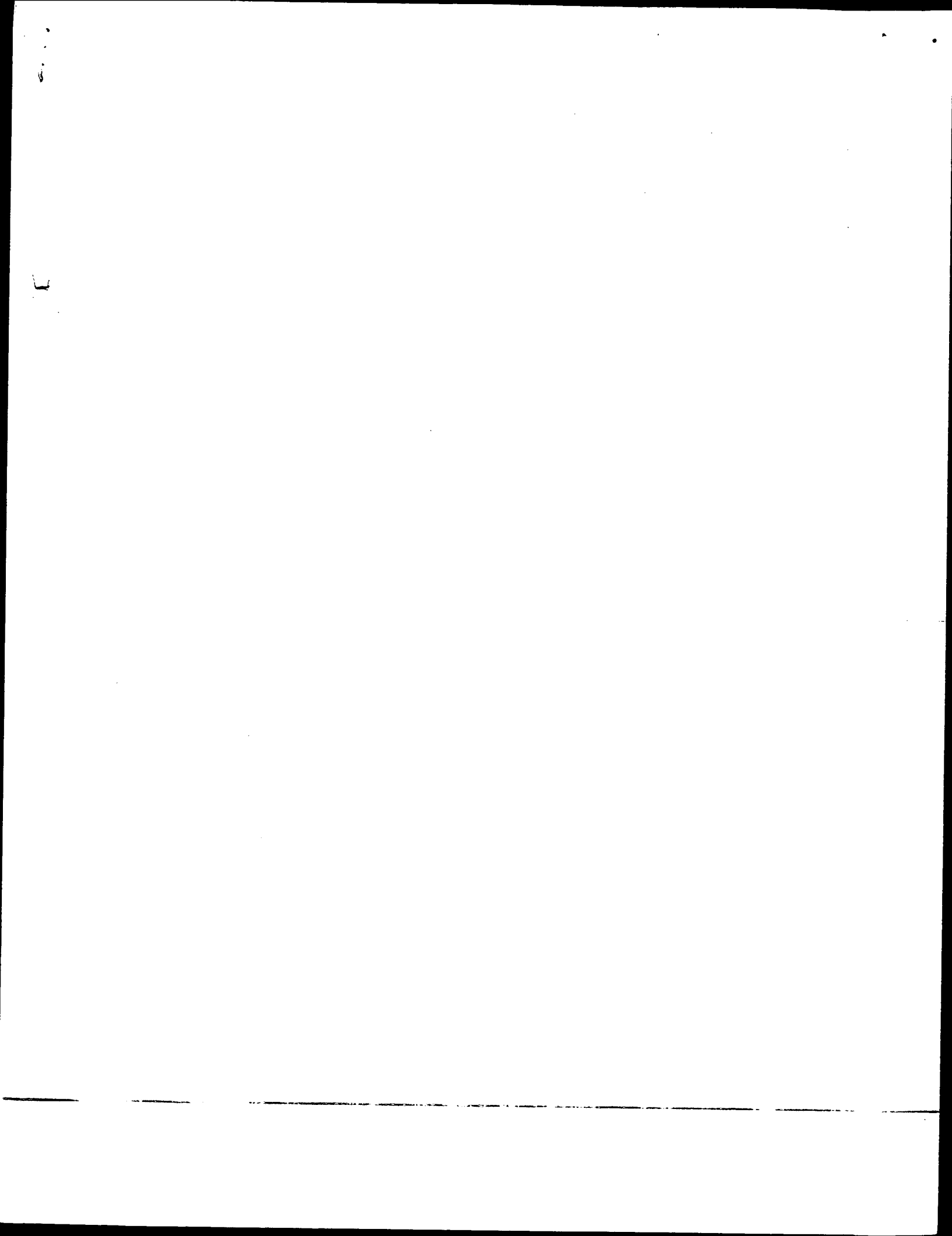
? APPLICANT: Wisniewski, Nancy
? APPLICANT: Silver, Gary M.
? APPLICANT: Lo, Katherine C.
? APPLICANT: Brandt, Kevin S.
? TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES
? TITLE OF INVENTION: PROTEINS AND USES THEREOF
? FILE REFERENCE: FC-3-C1
? CURRENT APPLICATION NUMBER: US/09/182,816
? CURRENT FILING DATE: 1998-10-29
? EARLIER APPLICATION NUMBER: 08/989,510
? EARLIER FILING DATE: 1997-12-12
? NUMBER OF SEQ ID NOS: 31
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 22
? LENGTH: 1736
? TYPE: DNA
? ORGANISM: Ctenocephalides felis
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (159)..(1553)
? US-09-182-816-22

```

	Query Match	1.3%	Score 39.6	DB 4	Length 1736
	Best Local Similarity	61.8%	Pred. No. 0.27		
	Matches	63	Conservative	0	Mismatches 39
					Indels 0
					Gaps 0
QY	2844	ttgtactcttttaagcgaataatcattttgttttgcctttccatttcctt	2903		
Db	1727	TT	1668		
QY	2904	ttgtcaaatatataataagattcattcaaatagattacc	2945		
Db	1667	TTTTTTTTTTTATATTGACAAGTTGAAATATAAATCCATATC	1626		

Thu Jun 28 10:23:27 2001

us-09-541-094-13.rni



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 27, 2001, 14:35:06 ; Search time 2158.27 Seconds
(without alignments)
12916.116 Million cell updates/sec

Title: US-09-541-094-13

Perfect score: 2949

Sequence: 1 tctgcagaatcggcttgcg.....atcaaatagatataccac 2949

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

JOURNAL
COMMENT

Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuko, Y., Isogai, T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3' - end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

```

1.840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP4001941"
/clone_id="NT2RP4"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME16SEFL3; mRNA from NT2 neuronal precursor cells after 2 weeks retinoic acid (RA) induction"

```

BASE COUN
ORIGIN

Query Match	28.0%;	Score 827;	DB 108;	Length 840;
Best Local Similarity	99.0%;	Pred. No. 3.8e-212;		
Matches 830;	Conservative	0;	Mismatches 8;	Indels 0;
				Gaps 0

Query 1

1 Similarity	99.0%;	Pred. No. 3.8e-212;		
830; Conservative	0;	Mismatches	8;	Indels
				0;
				Gaps
				0.

QY 121

AGCAGAGAGCGAATGCTACGGCAGGGGTGCTCTGGGGCTGACCCGGAAATCG 62

1800

122

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918

DB 12

gagcaactatcagattgctgcagctctcaattagtgtgagacacaggggttatccacgtta 367

Db 18

gttagagaagagagacctaagtgtgtatctgactgtgccccacccttacatg 427

Db 24

ttctactggaagcaacatttaccaggatttaatggaagctgaaagggaacc 487

25

547

02

AGCCGAATTCGTGCTTGCAGTGTCTTGACCAAGCCAGTCTGCCCCAGGCTTCTC 422

QY 54

CCTAGTGTACAGTGGCCAAATGATGGGTTTGGTGTCTACTCCAATTCCCTATGGCCAGAG 482

QY 60

TTTGCTCACTGCAGAGAAATACAGTGGAAATTCGCTGGGCATGCTTTGGCTTATGAAGAC 542

2

502

1

tatcaagatcacaaacctgagtcagaagtgtctcagcaccacacccctccacatratggtgcacg /8/

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	827	28.0	840	108	AU133372	AU133372	AU133372 AU133372
2	797	27.0	853	154	BG489641	BG489641	BG489641 602518064
3	768.2	26.0	838	108	AU132380	AU132380	AU132380 AU132380
4	753.8	25.6	784	19	A137767	A137767	A137767 fe62b05.x
5	691.4	23.04	748	165	BE262466	BE262466	BE262466 601154772
6	677.8	23.4	914	147	BF339787	BF339787	BF339787 602034987
7	666.2	22.6	696	155	BG568503	BG568503	BG568503 602587478
8	661.4	22.4	910	147	BF340237	BF340237	BF340237 602036827
9	652.8	22.1	807	167	BE408912	BE408912	BE408912 601303794
10	634.8	21.5	950	172	BF880176	BF880176	BF880176 602288192
11	633.2	21.5	656	169	BF793556	BF793556	BF793556 602255077
12	628.6	21.3	933	150	BF370121	BF370121	BF370121 602185971
13	627	21.3	934	150	BF530600	BF530600	BF530600 602071875
14	623	21.1	813	150	BE569485	BE569485	BE569485 602185971
15	606.4	20.6	820	140	BE788789	BE788789	BE788789 601475883
16	597.2	20.3	622	143	BF035070	BF035070	BF035070 601456171
17	596.8	20.3	622	143	BF979552	BF979552	BF979552 602288192
18	593	20.1	920	172	AU125659	AU125659	AU125659 AU125659
19	593	20.1	779	107	AU125659	AU125659	AU125659 602424205
20	591.6	20.1	808	152	BG325507	BG325507	BG325507 602424205
21	587.2	19.9	604	147	BF375675	BF375675	BF375675 RC5-ST030
22	585.8	19.8	678	168	BE680467	BE680467	BE680467 602154273
23	582.4	19.7	639	164	BE158319	BE158319	BE158319 RC2-HT038
24	580.2	19.7	639	115	AM372837	AM372837	AM372837 RC1-BT037
25	580.2	19.7	916	141	BE507949	BE507949	BE507949 601497312
26	576.4	19.5	618	115	AM386734	AM386734	AM386734 RC0-PT002
27	574.4	19.5	602	115	AM386847	AM386847	AM386847 CM4-HT019
28	572.4	19.4	913	141	BE309665	BE309665	BE309665 601501673
29	572.4	19.4	598	143	BF035344	BF035344	BF035344 601485511
30	570.4	19.3	584	141	BE890486	BE890486	BE890486 601431556
31	565.8	19.2	670	167	BE388123	BE388123	BE388123 601284450
32	561.8	19.1	610	152	BG325127	BG325127	BG325127 602423676
33	560.2	19.0	655	32	AV699416	AV699416	AV699416 AV699416
34	558.2	18.9	627	107	AU128443	AU128443	AU128443 AU128443
35	558.2	18.9	596	105	AL039414	AL039414	AL039414 DKFZ43414
36	555	18.9	600	115	AM369337	AM369337	AM369337 OV1-BN000
37	548	18.6	604	107	AU127220	AU127220	AU127220 AU127220
38	544.8	18.5	602	141	BE878476	BE878476	BE878476 601487888
39	537.8	18.0	546	125	AM805004	AM805004	AM805004 OVA-UM009
40	531.6	17.9	554	110	AM363805	AM363805	AM363805 RC0-CT035
41	525.4	17.8	558	164	BE144286	BE144286	BE144286 MR0-HT016
42	524.4	17.8	528	159	BE709658	BE709658	BE709658 I13-HT064
43	520.6	17.7	565	164	BE144270	BE144270	BE144270 MR0-HT016
44	520.6	17.7	701	150	BF537398	BF537398	BF537398 602050035
45	520.4	17.6	738	140	BE787574	BE787574	BE787574 6014794653

ALIGNMENTS

AU128443	AU128443
AL039414	DFP2343L
AW369337	VI-BN000
AU127220	AU127220
BE878476	6014872888
AW850504	QV4-UM009
AW366805	RCO-CT0355
BE144266	MRO-HT016
BE709658	IL3-HT061
BE144270	MKO-HT016
BE537398	602050035
BE781574	601479463

RESULT	1
AU133372	
LOCUS	
DEFINITION	AU133372 840 bp mRNA EST 24-OCT-2000
ACCESSION	AU133372 NT2RP4 Homo sapiens cDNA clone NT2RP4001941 5', mRNA sequence.
VERSION	AU133372
KEYWORDS	AU133372.1 GI:10993911
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Negai,T., Sugano,S., Masuno,Y. and Isega,T.
TITLE	HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,

	Dy	603	TATCAAGATACAAACTGAGTGCAGATGGCTCAGACCACCTTTCCACTAATGTGCATG	662
	OY	788	cagccttttcaacatgcattgctgttaacagcacctgccacctgcatlgcgcgcaagtcc	847
	Dd	663	CAGCTCTTTTCACACATGCATGCTGTCTCATCGACCTGCCACTGCATMGCGGCACCTTC	722
	OY	848	atccaagaaccttaagcatcaaccaggaaatgtctgttaaccocctgtctgaataaat	907
	Dd	723	ATCCAAGAAGACCTTGAGCATCACCCAGAAATGCTGTGTGAACCCCTGCTGTGAATACAT	782
	OY	908	gtgtggagcatgctaaagacctataataacaacttggcgacattaagcctagcagcaagg	965
	Dd	783	GTCGTGAGCATGCTTAAGCCTATAAATACACACTGGACATTNAAGCTGCANACANGG	840
	RESULT	2		
	LOCUS	BG489641	853 bp mRNA EST	27-MAR-2001
	DEFINITION	60251806/f1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:463670 5',		
	ACCESSION	BG489641		
	VERSION	BG489641.1 GI:13451151		
	KEYWORDS	EST.		
	SOURCE	human.		
	ORGANISM	Homo sapiens		
	REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	AUTHORS	Eumariyola; Eutheria; Primates; Carnivora; Insectivora; Homioidea; Homo.		
	TITLE	NIH-MGC http://mgs.nci.nih.gov/		
	JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
	COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapds-femail.nih.gov Tissue Procurement: DCTD/DTP/Gazdar CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LINC1395 row: P column: 15 High quality sequence stop: 853.		
	FEATURES	Location/Qualifiers		
	SOURCE	1..853		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:463670"		
		/clone_lib="NIH_MGC_18"		
		/tissue_type="large cell carcinoma"		
		/lab_host="DH10B (phage-resistant)"		
		/note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2:		
		EcoRI; CDNA made by oligo-dT priming. Directionally clone		
		into EcoRI/XhoI sites using the following 5' adaptor:		
		GGGACGAC(G). Library constructed by Ling Hong in the		
		Laboratory of Gerald M. Rubin (University of California,		
		Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and		
		Superscript II RT (Life Technologies). Note: this is a		
		NIH_MGC Library."		
	BASE COUNT	202 a 237 c 211 g 203 t		
	ORIGIN			
	Query Match	27.0%; Score 797; DB 154; Length 853;		
	Best Local Similarity	97.4%; Pied No. 4.8e-204;		
	Matches	831; Conservative 0; Mismatches 20; Indels 2; Gaps 2		
	OY	1324	ggagagatcttccttgccacattggagagatggtgtgtgtgttccctgctgcatcctcag	1383
	Dd	2	GGAAGATCTCTCTGGCACATTGGAGAAAGTGTGCTGTGCTCCCTGCTGCATCTCCAG	61
	OY	1384	ggagccaatatcagttcccagcctctccacacatcttccctgacgagattcttcgagctg	1443

Db	62	GAGGCAAAATGACGTCACGGCTCTCCACCAATCTTCCGACGACGATTTCTTCAGACTCG	121
Qy	1444	aaacatctctggcgtgtgtcttgcgtcyacacatctggtgcttccatacaaatatlaacp	1503
Db	122	AAACATCTCTGGCCCTTTCTTGCTGCTACCACTCTGCTGCTTCATCAAAATATTTACCA	181
Qy	1504	gagatattaacacacatctctgagagacattatgtgagatcccgaaatgycgtgagccctg	1563
Db	182	CAGATTATTAGACACTGCTAGAACACTTATATGATGAGCTATCCGAAATGGCTGACGACCTGA	241
Qy	1564	agaggaacctgaacttlytlaacagacacatcyccaagggcccttggacatlytggccaagtgct	1623
Db	242	AGAGGACCTGAACCTTTGTATACACACACTCTCCAAAGGCCCTGGAGATGGGACACGGTGCT	301
Qy	1624	gggagcgtgctctgtatgagagcttgcgagaggaacacattcagacagacagttcaagctga	1683
Db	302	GGGAGCGTCTGTATGATGAGCTGTCCAGAGGAACCAATTCAGCACACAGTTCAAGCTGA	361
Qy	1684	tcaccaaacagttatcccgccctgtctatctatgtgttccctgtataagccaacaactcatggt	1743
Db	362	TCGCCAAACGGTTACCCGCGCTGCTGTATGGGTTCCTGATTAACCAACACTCATGTGTT	421
Qy	1744	ccagttcatctccatcaggcagaggaacctaaagttcctacttctgttgggtgaaggcccttcaacat	1803
Db	422	CCAGTCTATCTCTCAGGACAGACACTTAAGGCTCACTTGGGTGACGCGGCTCTTCACAAATTA	481
Qy	1804	catgcgtctctcagcccccaacaacacacattatgttataagatctcttggcaatt	1863
Db	482	CATGCGTGTCTCCAGGCCCAACCAACCACTTAATGTTTACAGATATGCTTGGCAAAATTT	541
Qy	1864	gacttgcacagttgttcaacctcaccgcagagagagtgccagatgccaaatlaagttcccaag	1923
Db	542	GACTGGCACAGTGTGTCACCTCAACCCGAGAGAGAGTCCAGGATTCAGTAAGTCCCAAG	601
Qy	1924	tgaanaacagatctctatgatatctatctatgtgtccagggcccttgcattctaaatgaagc	1983
Db	602	TGAANAACAGATCTCTATGATGATCTCAATGAGTCTCCAGGCGCTTTGCATTTCAATGAGAC	661
Qy	1984	ggacgacatcccccgcggtgtgtgtgtcttctactctgcagcatlaagccagggccctgtctctc	2043
Db	662	GGACCGACTCCCCGGGTGTGTGCG-TCTATGACAGATTTAGCCAGATGCTTGTCTCTCTCG	720
Qy	2044	ctttgacatcagttatcagttgagagctctactcgaatctactacatgatcagatgcgtgtgaa	2103
Db	721	CTTTGACATCAGATGATGAGTGTGATCTTACATGATTTCTTACATGACATGAGCGCTGGAAA	780
Qy	2104	agataatcgttccgcgatatattctcatcgcacgacaagaagcttggatgtatcacctgt-2	2162
Db	781	AGGATATCCGTGGCGGATTTTCTCATGTGCAAGCAAAAGACGTGATGTATCACCTTGAA	840
Qy	2163	cagttgggcttgcg 2175	
Db	841	CAGTGGGCTTCGG 853	
RESULT 3			
AUI12380	838 bp	mrna	EST 24-OCT-2000
LOCUS			
DEFINITION	AUI12380 NT2RP3 Homo sapiens cdna clone NT2RP3004312 5', mRNA		
ACCESSION	AUI12380		
VERSION	AUI12380.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo;		
	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,		
	Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and		
	Isogai,T.		
TITLE	HRI human cdna project		
JOURNAL	Unpublished (2000)		

COMMENT

Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

Location/Qualifiers
1. 838

BASE COUNT 221 a 207 c 202 g 206 t 2 others
ORIGIN
Query Match 26.0% Score 768.2; DB 108; Length 838;
Best Local Similarity 98.8%; Pred. No. 2.9e-196;
Matches 815; Conservative 0; Mismatches 5; Indels 5; Gaps 4;

132 agagaggaagaatgctacgagcagggggtgctctggtggtgacccggggaagctggtc 191
17 AGAAGAGGCAAGATGCTACGACGAGGGGTGCTGTGGGGCTGACCCGGAGGTCGGGGTC 76
192 tcttcgct 251
77 TCCTTCGCT 136
252 agaggaagataatatacccttaataaagacgtccctgtctgtcgtctctcaagcga 311
137 AGAGGAGATATATATCCCTTAATAAACAGCTCCCTGTCTGCTCTCAACGCCA 196
312 ctcaatcagatggtcgtcagcgtctcaatctagtgagacagaggggttaaccagtag 371
197 CTCATCAGATGGGTGTCAGGCTCTCAATTAAGTGAGACACAGGGGTATCCACGTAG 256
372 agaagagagagacacagctggtggtggtggtggtggtggtggtggtggtggtggtc 431
257 AGAAGAGGAGGAGACTACAGTGGTATGACTGATGAGCCCAACCCCTTACATGATT 316
432 tgcctgagagagacacattaccagaggttaataagagagagagagagagagagagc 491
317 TCCTGGAGAGCAAGCATTTTACAGGGATTAATGAGAAAGCTGAAGGAGAAACCAAC 376
492 gaattggtggtcttgcagagctctgacacagccagctcctcagcgtctctccta 551
377 GAATGCTGTGCTTGCAGGTCTTGCACCAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 436
552 gttacagtgcccaaatgattggtgtgttactccaattccattatgagcagagtttg 611
437 GGTACAGTGCACCAATGATGAGTTGGTTTACTCCAAATTCATGAGGACAGATTG 496
612 ctcaatcagagagaataacagtggaatgctggtggaatggttggcttaagaaactta 671
497 CTCACAGCAGAAATACAGTGAATTCCTGGGCAATGGTGGCTTATGAGAACTTTA 556
672 gttcccatctctctctctgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 731
557 GTTCCCATCTTCTTCTTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 616
732 aagatcaaacctgagtcagaaatgctcagcaccacacacacacacacacacacacacac 791
617 AAGATCAGCACTGAGTCAAGATGCTCAGCAGCAACACTT-CCACTATGTGCAATGACG 675

QY 792 tctttcacatgcatgctctctacacagcactgcccactgcatcggcagcagctcatcc 851
DB 676 TCTTTACACATGATGCTGCTGATCAGCACTGACCTGCATCGGGGCA-CTTCAATCC 734
QY 852 aagacacttgaacataaccagaatcgtctgacccctctctgattacatgct 911
DB 735 AAGCACTCTGACATCAACCAAGAAATCGTGTGACCCCTGTCATTAAGATGTGT 794
QY 912 -ggagatgctaaagcctataataacactggagacataagcct 955
DB 795 GGCAGCATCTTAAGCTATATA-TNCACTGGGACATTNAGCCT 837

RESULT 4

AI377767 784 bp mRNA EST 28-MAR-1999
LOCUS t662b05.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:2091249 3' similar to FR:Q92542 MYELOBLAST KIAA0253 ;
mRNA sequence.

ACCESSION AI377767 GI:4187620
VERSION AI377767.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 784)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through INLIL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 450 Std Error: 0.00
Seq primer: -40UP from g1bco
High quality sequence stop: 462.

FEATURES

Location/Qualifiers

1. 784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19w, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 238 a 141 c 232 g 172 t 1 others
ORIGIN

Query Match 25.6% Score 753.8; DB 19; Length 784;
Best Local Similarity 98.3%; Pred. No. 2.2e-192;
Matches 769; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

QY 2168 agcttcggcaccatctatcttcctcctacatgctacactgcatcaatgccaagctgat 2227
DB 784 GGGTTCGGCATCTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 725
QY 2228 gtcccttcattgctcccgaggagcagagctgtgtctactatgagagagcscagct 2287
DB 724 GCTTTTTCATTTGCTCCCGGAGGACGAGACTGTGCTATCTGAGGAGGACCCACACTT 665

DB 84 CCGATTACACCCCTAA-TTGCCCTTCAGAGACCTTCTACATTTCCGCCGCCCTCCG 26
 QY 2615 accctctctgtctccacc 2635
 DB 25 CCGATTACACCCGCTGACC 5

RESULT 6
 BF339787 914 bp mRNA EST 22-NOV-2000
 LOCUS 602034967F1 NCI_GCAP_Brn64 Homo sapiens cDNA clone IMAGE:4183101
 DEFINITION 5', mRNA sequence.
 ACCESSION BF339787
 VERSION BF339787.1 GI:11286246
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 914)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: cgabs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM9498 row: m column: 22
 High quality sequence stop: 670.
 Location/Qualifiers

FEATURES
 source 1..914
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="NCI_GCAP_Brn64"
 /library="gliblastoma with EGFR amplification"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: brain; Vector: PCMV-SPOrt6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 Kb. Constructed by Life Technologies. Note: this is a NCI_GCAP Library."

BASE COUNT 197 a 268 c 226 g 223 t

ORIGIN

Query Match 23.0%, Score 677.8; DB 147; Length 914;
 Best Local Similarity 97.0%; Pred. NO. 8e-172;
 Matches 752; Conservative 2; Mismatches 14; Indels 7; Gaps 6;

QY 1570 cctgaacttgaagaagcactgcaagcctgacagatgagcagctgctgagcagc 1629
 DB 1 CCGTAACCTTTGACAGACACTGCCAAGGCTCCGCAAGTGTGGCCACGAGGCTGGAGC 60

QY 1630 tgcctgtatgagcttgagaggaaccaacttcagcgacagcttcagctgaccca 1689
 DB 61 TCGCTGATGAGCTTGCGAGGAGAACCAACTTCAGCGACAGTTCAGGGCTGATCCCCA 120

QY 1690 aacgggtaccgcctgctctatgggtctctgattaaagcaacaactatggtccagtc 1749
 DB 121 AACGGTTACCCCTGCTTATGGGTTCTGATTAAGCAACAACACTCATGCTCCAGTC 180

QY 1750 tatctcagcaggaactaaggctcacttggtgtagcagggctcttcaacattacacgc 1809
 DB 181 TATCTCAGGACGACCTAAGCTTACTGGGTGACGGGCTCTTCACATTAATACGC 240

QY 1810 tgcctcagcccaacaacaactatgttgcacagctgctctggcaattgacgag 1869
 DB 241 TGTCTCAGCCGCCACCAACACACTTATGTGTACAGTATGCTTGCAAAATTTGACTGG 300

QY 1870 cacaagtcaacctaccacagagcagctgcacagatccaaagtaagtcccaagtgaab 1929
 DB 301 CACAGTGTCAACCTCACCCAGAGCACTGCCAGATCCAAAGTCCCAAGTGA 360

QY 1930 caagatctgtatgagtaactcagagcagcagccttgatcctaataagagagcagc 1989
 DB 361 CAGAGATCTGTAGAGTACTCATGAGGTCCA-GGCCCTTGCATCTAATAGAGGAGCC 419

QY 1990 actcccccgtgtgctgcttactcagatagcagagcctgtctcctcttga 2049
 DB 420 ACTCCCCGGTGTGCGTCTACTGACAGATTACCGAGGCTTGTCTCCTGCTTGA 479

QY 2050 actgagtcagtgagcttactcagatcactcagagcagcagcctgaaagat 2109
 DB 480 ACTGAGTCAGTGAGCTCTACTGAATCTCATAGCATGAGAGAGCCCTGGAAAGAT 539

QY 2110 ccgtgccgatatctcctcagcagcaagagcttgatgataccctgacagtgag 2169
 DB 540 CCGTGCCCGGATATTTCATCGCCAGCAAGAGCTTGATGATCACCCTGACAGTG 599

QY 2170 ctgcagcactcactcttctcctcactgacactcagatcactcagatcagctgag 2229
 DB 600 CTTCGCGATCTCATCTCTCTCCCTCATCTGCTCATCTG-ATCAATGCCAAGCTGAT 658

QY 2230 ccttctcattgctcccccggag--ccagagagctgtgcatc-tgagagagccagct 2286
 DB 659 CCTTTCATTTCTCCCGGAGGAGCCAGAGGCTGGTCTACTTGGGAGGAACCAAGCT 718

QY 2287 ttctctcagctc-agcagctcact-ctcagagcactgtcccactggagcagc 2339
 DB 719 TTCTTCGCGAGCTCAAGCAAGTTCATCTCCCTAGAGCATTTGCCCTCGGACAC 773

RESULT 7
 BG568503 696 bp mRNA EST 10-APR-2001
 LOCUS 602587478F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:471626 5',
 DEFINITION mRNA sequence.
 ACCESSION BG568503
 VERSION BG568503.1 GI:13576156
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 696)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: cgabs-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM1564 row: k column: 19
 High quality sequence stop: 696.
 Location/Qualifiers

FEATURES
 source 1..696
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="IMAGE:471626"
 /clone_lib="NIH_MGC_76"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site: 1: SfiI (ggccgctggcc); Site: 2: SfiI (ggccatcagagc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATATGCGC-3' and 3' adaptor sequence: 5'-ATTGTAGAGCGGAGCGGCGGACATG-dt(30)BN-3' (where B = A,


```

|||||
DB 541 CAGGGCAACATACATGAGCCCTAGACAGGGGACATCTGGGCTAG-CTACTGTC 599
|||
QY 2751 tccctccacgtcctcttcacagccctcagatgacattagggtggcgtcgcg 2810
|||||
DB 600 TCCCTCCCACTGCTCTTCTCCAGG-CTCAGATGACACATTA-GGTGGGCTGCTGCG 657
|||
QY 2811 gggggtacccacccacccacagtgctcagtgtaatttatttaagcgttaattc 2870
|||||
DB 658 GTGGGTAT-CCACCTCCACCCACACAGTGTCAGT-GTACTTTTATTAAGCTGATATATC 714
|||||
QY 2871 taatttgcttcttcttcttcttcttcttcttcttcttcttcttcttcttca 2930
|||||
DB 715 TATT-----TGGTTGGCTTTTCTTCTTAACTTGTACTATCTTTTTCACAGGTTCTCA 770
|||||

RESULT 9
BEA08912 807 bp mRNA EST 21-JUL-2000
LOCUS 601303794P1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638182 5',
DEFINITION mRNA sequence.
ACCESSION BEA08912
VERSION BEA08912.1 GI:9345362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 807)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM338 row: 1 column: 23
High quality sequence stop: 610.
Location/Qualifiers
1. 807
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3638182"
/clone.lib="NIH_MGC_21"
/tissue.type="choriocarcinoma"
/lab_host="PH10B (phage-resistant)"
/note="Organ: Placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(s). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 167 a 211 c 220 g 209 t
ORIGIN
Query Match 22.1%; Score 652.8; DB 167; Length 807;
Best Local Similarity 95.8%; Pred. No. 4.4e-165;
Matches 692; Conservative 0; Mismatches 27; Indels 3; Gaps 2;

QY 938 acgggacataaagcctgacagaggtgtgtgtcctccacccgctgatacgt 997
|||||
DB 1 ACTGGGACATTAAGGCTGACGACAGGGTGTGCTTCTCTCCACCGGCTGATGTCGT 60
|||||
QY 998 tcccttcttgtaatgtgcccacaggggctgaagcgcagtgcttcttcttccaccag 1057
|||||
DB 61 TCCCTTTTCTGGAATGTGGCCCGGAGGCTGAAGGCGCAGTGTCTTCTTGTACCCAG 120
|||||
```

```

QY 1058 ctggcctgtctgtaagcttgcacaaagccacttgatgtgaccaccctcccgcaatgtc 1117
|||||
DB 121 CTGGCTGTCTGTGAAGCTTTTCAAAAGGACCTTGATGTACACCTCTGCCCAATGTC 180
|||||
QY 1118 atgttgcctcttcttcaagggaacttcttgaactatctgagcagctgagatgtctac 1177
|||||
DB 181 ATGTTGTCTTCTTCAAGGGGAAACTTTTCACTACATTTGGCAGCTCGAGGATGTCTAC 240
|||||
QY 1178 gatatgaaagggaagcttcccgctcagctagaagaatgttgaactatctgtgagctg 1237
|||||
DB 241 GATATGAGGAAGGGGCAAGTTTCCGTCAGTATAGAGATGTTGACTATTTGTGTGAGACTS 300
|||||
QY 1238 ggaacagtgagccttaagaaacttcaatagagcttggatggacaaagatccgtttccag 1297
|||||
DB 301 GGACAGGTGGCTTTAAGAACTTCAATTAAGCTTGTGATGCACACAGATCCTGTTTCTGAG 360
|||||
QY 1298 aaaaatgagctgtacaggaacaggtgagagatctcctgacacattgagaagatggt 1357
|||||
DB 361 AAAATGAGTCTGTACGAGACAGGTGAGATCTCCTGGCCACATTTGGAAGAGTGTG 420
|||||
QY 1358 gctggtctccctgctgtcatcctcgaagagggccaaatcagttccagcctctccacatc 1417
|||||
DB 421 GCTGTGTCTCTGCTGTCTATCTCTCAGAGCCAAATAGTCCAGCCTCTCCACATCT 480
|||||
QY 1418 tccctgcagcagatcttctcgcagctcgaacacatctcgtgctgtctgtgctgaccactc 1477
|||||
DB 481 TCCCTGCAGCAGATTTCTTGTGAGCTGGAACATCTCTGCTGTGTGTGCTGACCACTCT 540
|||||
QY 1478 gctgctccataacaaatattaccagagatttaccagacatctgtgagaacattatgt 1537
|||||
DB 541 GGTGCTTCCATTAACAATATTACAGAGTATTACAGACATCTGCTGGAACATTAAATG 600
|||||
QY 1538 agctatccggaatgctgagccct--gaagagagactgaacttgttaacg-acactgctc 1554
|||||
DB 601 AGCTATCCCGAAATGTTGAGAGCTTGAAGAGGAGTGTATCTTGTATACAAACCTTGC 660
|||||
QY 1595 aagggccctggcagatgtgtgccaaggtgtgagcgtgtgtgtgtgtgtgtgtgtgtgt 1654
|||||
DB 661 AAGGCCCTGGCAATGTGTGGCCCGGGGCTGGGAGCTGCTGTGTGTGTGTGTGTGTGTGT 720
|||||
QY 1655 ac 1656
DB 721 CC 722

RESULT 10
BE980176 950 bp mRNA EST 23-JAN-2001
LOCUS BE980176/6
DEFINITION 60228192P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:437381 3',
DEFINITION mRNA sequence.
ACCESSION BE980176
VERSION BE980176.1 GI:12347391
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 950)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10036 row: h column: 24
```

FEATURES	
High quality sequence start: 24	
High quality sequence stop: 729	
Location/Qualifiers	
1. .950	
source	

BASE COUNT ORIGIN	282 a	202 c	280 g	186 t
----------------------	-------	-------	-------	-------

Query Match	21.5%;	Score 634.8;	DB 172;	Length 950;
Best Local Similarity	92.9%;	Pred. No. 3.4e-160;		
Matches 784;	Conservative	2;	Mismatches 44;	Indels 14; Gaps 11

QY	2058	agtgagagctctacgaaatactcacaatgacatgagccgcctggaagaatcgc-gtcc	2118
Db	853	AGTGTGGGTTGTGTACTACAGTGAAGTGAATCTCTGGAGAGATATCCGGGGCC	794
QY	2117	cggatattctcatcgcgcagcaaaagccttgag---tgaataccctgacg-tggcctt	2172
Db	793	CGGATATTTCTTATTCGAAACCAAGAACCGTGAGGTGATATCCGTGACATTTGGGTTT	734
QY	2173	cggatcccatctcttccctcactgctcaact-acgcaatcaatgccaaagtat-gt	2229
Db	733	CGGATCCCTATCTCTCCGTATCGTCAAGTACATGCATCAATGCCAAAGTGAATTGTC	674
QY	2230	cccttctatgtctcccggaagccagagcgtgltcaatactgagaaagccacagcttt	2289
Db	673	CGTTTTCATTGCTCCCGGAGGACAGAGACTGTGTCATCTAGAGAGACCCTCACCTTTT	614
QY	2290	-cttgacagctcgaagatlaacttctctagagatctgctccactgaggacaacaacaa	2348
Db	613	CTTTGCCAGCTCAGCACTTACTTCTCTTAAGCATCTGTCTCCACTGGACACACCATGAA	554
QY	2349	ttgtgcactggaacctccctcctgagctcagattggatgaataaagaagtgagac	2408
Db	553	TTTGTCTACTGGAAACCT-CCGAGGCGCTGTCTCAGATTGGGATTAACTAAAGAGTGAAAC	495
QY	2409	tattccaaagaagacaaggga-gaaataataaatgtctccctctctccgctcccttcc	2467
Db	494	TATTCATTAAGAGACAGGAGATTAATAAATTAATGTCTCCCTCCGCTCCCTCTTCC	435
QY	2468	catcacccctctcccaattctctctctctctactcaatcagacagatttggatatacaa	2527
Db	434	CATCACCCCTCCGCCATTTCTCTTCTCTCTCTACTCAATGCCAGATTTGGGATTACAA	375
QY	2528	tagaagctctgctcctcgtttaactcctagtaaccaaccttaattgcccctoaagac	2587
Db	374	TAAAGGCTTCTCTCTCTTTAACTCCCTAATTACCAACCTTAATTTGGCCTTCAGAC	315
QY	2588	cctcttaactttctctctcctgacctgtactctctctcctctacaccccaacctgac	2647
Db	314	CTTTCTACTTTTCTCTCTCCGCTGTACTCTCTCTGTCTGTCTGCTCTCACCCCAACCCCTGAC	255
QY	2648	ccagcaacctctctgactggaagagacataaaagglttaatgctcagggctcaactacat	2707
Db	254	CCAGCCACCTCTCTGACTGGGAAGGACATAAAGGTTTAATGTACGGGTCAACACTT	195
QY	2708	gagccc-ctagagagacaggacatctctctgagcgaactacgtctcctctcccaatgctc	2766
Db	194	GAGCCCTCTGAGGACAGGGGCAATCTCTGGGCTGAGCACTGATCTCTCTCCACTGTGCT	135

QY	2767	ttctccagagcccttaagttggcaaatagggttggtctgctg-ccagttggtatcccaact	2825
Db	134	TTTCACAGGCCCTCAGTAGGCACATTNAGGCTGGCGCTGCAGCGGATGCCACACT	75
QY	2826	ccagccacagctgctcagttgtaacttttaataagctgaataatcatttgcgttttgg	2885
Db	74	CCAGCCACACAGTGTCTAGTTGTAACCTTTTA-TAAGCTGTAAATATCTATTGTTTATATCA	16
QY	2886	cttt 2889	
Db	15	CTTT 12	

RESULT	11
LOCUS	BF793556
DEFINITION	BF793556 656 bp mRNA EST 12-JAN-2001 602825077.F1 NIH_MGC_84 Homo sapiens CDNA clone IMAGE:4347355
ACCESSION	mRNA sequence.
VERSION	BF793556
KEYWORDS	BF793556.1 GI:12098610
SOURCE	EST.
ORGANISM	human.

BASE COUNT	148 a	208 c	120 g	180 t	
ORIGIN					
Query Match		21.5%,	Score 633.2;	DB 169;	Length 656;
Best Local Similarity		99.4%;	Pred. No. 8.1e-160;		
Matches 654;		Conservative 2;	Mismatches 0;	Indels 2;	Gaps 2;
QY 2051	ctgagtcagctgagctcctcactgaataactcttaatgactgtagagccgcctggaatatatc	2110			
Db 1	CTGAGTCAGTGGAGCTCTACTGTAATACCTCTCATGAGACTGAAGGCCGCTGGGAAAGATATTC	60			
QY 2111	cgctcccgagatatcttcctatcgccagcaagaagcttgtagtgcataccctgacagtgggc	2170			
Db 61	CGTCCCGGATATTTCTCATCGCCAGCAAGAAGCTTGAGTTGATATCACTTCAGCATGGGC	120			
QY 2171	ctcgacatccatcatctcttcctccatcgctcaactactgcatlcaatgaccaagctgagtgc	2230			
Db 121	TTTCGGCATTCCTCATCTTCCTCATCGCTACCTCATCTGCATCAATGCCAAGCTGATGTC	180			

QY 2231 ctttctatctgctccgggagccaggagctgtctactaggaagaccagctttc 2290
 |||||||
 DB 181 cttttctgctcccccgggagccaggagctgtctactaggaagaccagctttt 240
 |||||||
 QY 2291 ttgcagactagagatctactctctagagcatctgtcccacttggagacacacactaat 2350
 |||||||
 DB 241 ttggcagactagagatctactctctagagcatctgtcccacttggagacacacactaat 300
 |||||||
 QY 2351 tgtacttggaacctccctgggctgtctcagattggatataaagaagtggaacta 2410
 |||||||
 DB 301 tgtacttggaacctccctgggctgtctcagattggatataaagaagtggaacta 359
 |||||||
 QY 2411 tcccaagagacagaggaataataatctgctccctccctccgctccctccat 2470
 |||||||
 DB 360 tcccaagagacagaggaataataatctgctccctccctccgctccctccat 419
 |||||||
 QY 2471 cacccttccctccctccctccctccctccctccctccctccctccctccat 2530
 |||||||
 DB 420 cacccttccctccctccctccctccctccctccctccctccctccctccat 478
 |||||||
 QY 2531 aagcttctgctcctgttactcctcctcctcctcctcctcctcctcctcctcct 2590
 |||||||
 DB 479 aagcttctgctcctgttactcctcctcctcctcctcctcctcctcctcctcct 538
 |||||||
 QY 2591 tctacttctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2650
 |||||||
 DB 539 tctacttctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 598
 |||||||
 QY 2651 gccaccttccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2708
 |||||||
 DB 599 gccaccttccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 656
 |||||||

RESULT 12
 BF570121/c 693 bp mRNA EST 12-DEC-2000
 LOCUS 60218597J171 NIH_MGC_45 Homo sapiens cDNA clone IMAGE:4310584 3',
 DEFINITION mRNA sequence.
 ACCESSION BF570121
 VERSION BF570121.1 GI:11643833
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 693)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Linehan
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNM at:
 http://image.llnl.gov
 Plate: ILNM185 row: m column: 17
 High quality sequence start: 27
 High quality sequence stop: 692.
 Location/Qualifiers
 1..693
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4310584"
 /clone_1ib="NIH_MGC_45"
 /tissue_type="renal carcinoma (ascites)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pORF7; Site:1; XhoI; Site:2;
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:

BASE COUNT 203 a 132 c 215 g 143 t
 ORIGIN
 GGCACGAG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library. "

Query Match 21.3%; Score 628.6; DB 150; Length 693;
 Best Local Similarity 97.2%; Pred. No. 1.4e-158;
 Matches 660; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 2228 gtccttctcattgctcccgag-----agccagagcctgtgtcactatcagagga-cgs 2281
 |||||||
 DB 679 gtctcttctcattgctcccgagccgggagccagagcctgtgttcaattcagaggaaccc 620
 |||||||
 QY 2282 cagcttctctcagagcctcagagatctactcctagagcatctgtcccacttggagacag 2341
 |||||||
 DB 619 cagcttctctcagagcctcagagatctactcctagagcatctgtcccacttggagacag 560
 |||||||
 QY 2342 ccactaatgtcacttggaacctccctgggctgtctcagattggatataaaga 2401
 |||||||
 DB 559 ccactaatgttgcacttggaacctccctgggctgtctcagattggatataaaga 500
 |||||||
 QY 2402 gtggaacttccaaagagagacagaggaataataatctgctccctccctccgctcc 2461
 |||||||
 DB 499 gtggaacttccaaagagagacagaggaataataatctgctccctccctccgctcc 440
 |||||||
 QY 2462 ctctcccatcacccttccctccctccctccctccctccctccctccctccctcc 2521
 |||||||
 DB 439 ctctcccatcacccttccctccctccctccctccctccctccctccctccctcc 380
 |||||||
 QY 2522 tacaataagaagcttctgtcctgttactcctcctcctcctcctcctcctcctcctcct 2581
 |||||||
 DB 379 tacaataagaagcttctgtcctgttactcctcctcctcctcctcctcctcctcctcct 320
 |||||||
 QY 2582 cagagcctctcacttcttctcctcctcctcctcctcctcctcctcctcctcctcct 2641
 |||||||
 DB 319 cagagcctctcacttcttctcctcctcctcctcctcctcctcctcctcctcctcct 260
 |||||||
 QY 2642 ctgtaccagacacactcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2701
 |||||||
 DB 259 ctgtaccagacacactcctcctcctcctcctcctcctcctcctcctcctcctcctcct 200
 |||||||
 QY 2702 tacttgaagccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2761
 |||||||
 DB 199 tacttgaagccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 140
 |||||||
 QY 2762 gtcccttctcagagcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2821
 |||||||
 DB 139 gtcccttctcagagcctcctcctcctcctcctcctcctcctcctcctcctcctcct 80
 |||||||
 QY 2822 accctccagccagcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2881
 |||||||
 DB 79 accctccagccagcctcctcctcctcctcctcctcctcctcctcctcctcctcct 20
 |||||||
 QY 2882 ttgtcttcttctcctatc 2900
 |||||||
 DB 19 ttgtcttcttctcctatc 1

RESULT 13
 BF530600 934 bp mRNA EST 11-DEC-2000
 LOCUS 602071875F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4214778
 DEFINITION 5' mRNA sequence.
 ACCESSION BF530600
 VERSION BF530600.1 GI:11617963
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 61 GGACCCAGCTTTCTGGCAGCTCAGACTTCTCTAGAGCTCTGCTCCACTGGG 120

Qy 2336 acgaaccactaattgtcaacttggaacccctccctggtcgtctcagattggaattacat 2395

Db 121 ACACAACCACTAA-TTGTCACTGAGACCTCCCTGGGCTGTCTCAGATTGGATTACAT 179

Qy 2396 aaaaagtggaactatccaaaagagcaaggaggaataataatgtcctccctcc 2455

Db 180 AAAAGGTGAGACTATCCAAAAGAGCAGGAGAAATTAATAA-TGCCTCCCTCTCC 238

Qy 2456 gctcccttcccatcaccctccctccctccctccctccctccctccctccctcc 2515

Db 239 GCTCCCTTCCCATCACCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 297

Qy 2516 tgggattaaataagaagcttctgctctgtttaaactccagtaaccacctaatt 2575

Db 298 TGGGATTACAAATAGAAAGCTTC-TGCTCTCTTTAACTCCCTAGTTACCCACCTAA-TT 355

Qy 2576 gccctcagaacctctactcttctcctcctcctcctcctcctcctcctcctcctc 2635

Db 356 GCCCTCAGACCTCTTCTACTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 415

Qy 2636 caaccctgtaccagcaacctcctcctcctcctcctcctcctcctcctcctcctc 2695

Db 416 CCACCCCTGTACCCAGCCACTTCTGACTGGAAGAGACATAAAGTTTATGTCAAGG 475

Qy 2696 taaactaatagaagccctcctcctcctcctcctcctcctcctcctcctcctcct 2755

Db 476 TCAAACTATATGAGCCCTGAGAGAGGAGGAGCTCTGGGCTGAGCCCTACTGCTCT 535

Qy 2756 cccactgtccttctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2815

Db 536 CCACCTGTCTCTTCTTCCAGGCGCTCAGATGAGACATTAAGGAGGCGCTGCGGCTGG 595

Qy 2816 tctccactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2875

Db 596 TATCCACCTCCAGCCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655

Qy 2876 tctgttctccttctccttctccttctccttctccttctccttctccttctccttct 2934

Db 656 T-GTTTGTCTTTTCTCTTATTC-TTTTGAATATATATATATATATATATATATAT 712

Qy 2935 aatagattatccc 2947

Db 713 AATTGAATTATCC 725

RESULT 15

LOCUS BE788789 820 bp mRNA EST 20-OCT-2000

DEFINITION BE788789 601475883F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3878780 5', mRNA sequence.

ACCESSION BE788789

VERSION BE788789.1 GI:10209976

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 820)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: DCTD/DRP/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1AM9643 row: e column: 21

High quality sequence stop: 698.

Location/Qualifiers

1-820

FEATURES

source

organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3878780"

/clone_id="NIH_MGC_68"

/issue_type="large cell carcinoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: Lung; Vector: pCMV-Sport6; Site-1: NotI; Site-2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 201 a 203 c 208 g 208 t

ORIGIN

Query Match 20.6%; Score 606.4; DB 140; Length 820;

Best Local Similarity 96.2%; Pred. No. 1.5e-152;

Matches 718; Conservative 0; Mismatches 16; Indels 12; Gaps 9;

Qy 449 ttaccagagatttaatgagaagctgaagagggaacacagccgaattgtgtcttgcga 508

Db 4 TTTTACCAAGGTTTAAATGAGAACTGAAAGGAGAAACAGCCGAATTTGCTGTTCGA 63

Qy 509 gtgtccttgacaaagccagctcctcctcctcctcctcctcctcctcctcctcctc 568

Db 64 GTGTCTTGACCAAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123

Qy 569 gatggttgggtgttctcctcctcctcctcctcctcctcctcctcctcctcctcct 628

Db 124 GATGGG-TTGGTGTACTCCAAATTCCTATGAGGCAAGTTTCTCTCACTGCAAGAAAT 182

Qy 629 cagtggaattcgtcggcgaatgtgtgtcctatgaagacttgaattcccatcttct 688

Db 183 CAGTGAATTCGCTGGGCAATAG-TTGGCTTATGAACACTTATGTTCCCATCTTTCT 241

Qy 689 ctgaagaatgaatgaacaaagctcaagcagtgctatcaagatcaacacctgaat 748

Db 242 CTGAAGATGAATAATGAACCAAGTCAACAGCTATCAAGATCAACAGTGAAG 301

Qy 749 caagatgtcctcagcaacaccttccactatgtgcatagtcagctctttcagacatgat 808

Db 302 CAGATGCTCAGACCAACCTTCCACTATGTGCCATGAGCTCTTTTCACATGACAT 361

Qy 809 gctgtcactcagcactgcacctgcatgctgctcagctcctcctcctcctcctcctc 868

Db 362 GCTGTATCAGCACTGCACCTGCATGCGCGGCGAGCTCCATCAAAAGCACTTACACATC 421

Qy 869 aaccagaatgtctgtgacccctcctcctcctcctcctcctcctcctcctcctcct 928

Db 422 AACCCAGAAATGCTGTGTGACCCCTGCTGATTAATGTGTGGAGCATGCTAAAGCT 481

Qy 929 ataatacaacttggaacatgaagcctgacagagaggttggtgtgtgtgtgtgtgtgt 988

Db 482 ATAAATCACTGGGACATTAAGCTGACGACAGAGGCTGTGTGTGTGTGTGTGTGTGT 541

Qy 989 gatagtgtccttcttctgaatgtgagcccaaggagctgaagcagctgtgtctcc-ct 1047

Db 542 GATAGTCGTTCTTTTCTGGAATGTGG-CCGAGGGGCTGAAGAGCCCAATGGCTTCTTT 600

Qy 1048 tctcaccacgtctgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1107

Db 601 TGTACCCACAGCTGCTGCTGCTGTAAGC-TTGGAAAAGGACCTGATGATGACCACTTG-C 658

Qy 1108 ccgcaatgtatgtgtgtccttctcctcctcctcctcctcctcctcctcctcctcctc 1167

Db 659 CCCCAATGTCATG-TGTCTCTCTTTTACGGGAA-----CTTGACTACTAGGGAGTCCGAG 713

Qy 1168 gatgtctacgatatgagaaggca 1193

Db 714 GAT-GTCTACGAATGAGGAAGGAA 738

Thu Jun 28 10:23:28 2001

us-09-541-094-13.rst

Search completed: June 27, 2001, 15:43:06
Job time: 4080 Sec

```

|||||
Db 2701 CTACATTGAGCCCTGAGGACAGGGGCTCTCTGGGCTGAGCCTACTGCTCTCTCCAC 2760
QY 2761 tgccttttccagccctcagatggacattaggggtggcgctgcgggtgggtatcc 2820
|||||
Db 2761 TGTCTTCTCCAGGCCCTCAGATGGCACCATTAGGTGGCGCTGCTGGGGTGGGTATCC 2820
QY 2821 cactccagccagctgctcagctgacttttatttaagctgaatactattttgtt 2880
|||||
Db 2821 CACCTCCAGCCACAGTGTCTGACGTGTTTATTAACTGCTATACTATTGTT 2880
QY 2881 ttgtcttttcttattctttttttaaataataataatagatttcaataaata 2940
|||||
Db 2881 TTTGCTTTTCTCTTATTCTTTTGTAAATATATATATATATATATATATATAT 2940
QY 2941 ttatccac 2949
Db 2941 TTATCCAC 2949

RESULT 2
LOCUS D87442
DEFINITION Human mRNA for KIAA0253 gene, partial cds.
ACCESSION D87442
VERSION D87442.1 GI:1665772
KEYWORDS KIAA0253.
SOURCE Homo sapiens male bone marrow myeloblast cell_line:KG-1 cDNA to mRNA, clone_lib:paluescript II SK clone:HA7036.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2805)
Nomura, N.
Direct Submission
Submitted (27-AUG-1996) to the DDBJ/EMBL/GenBank databases. Nobuo
Nomura, Kazusa DNA Research Institute, Gene Structure 1: 1532-3
Yana, Kisarazu, Chiba 292, Japan (E-mail:cdna@foekazusa.or.jp)
URL: http://www.kazusa.or.jp, Tel: 0438-52-3930, Fax: 0438-52-3931)
2 (sites)
Nagase, T., Seki, N., Ishikawa, K. and Nomura, N.
Prediction of the coding sequences of unidentified human genes. VI.
The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
analysis of cDNA clones from human cell line KG-1 and brain
Unpublished (1996)
3 (sites)
Nagase, T., Seki, N., Ishikawa, K., Ohira, M., Kawarabayashi, Y.,
Ohara, O., Tanaka, A., Kotani, H., Miyajima, N. and Nomura, N.
Prediction of the coding sequences of unidentified human genes. VI.
The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
analysis of cDNA clones from cell line KG-1 and brain
DNA Res. 3 (5), 321-329 (1996)
97191544
FEATURES
Location/Qualifiers
1..2805
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="KG-1"
/cell_type="myeloblast"
/clone="HA7036"
/clone_lib="pBluescript II SK"
/sex="male"
/tissue_type="bone marrow"
1..2128
/gene="KIAA0253"
<1..2128
/gene="KIAA0253"
/codon_start=2
/protein_id="BAA13383.1"
/db_xref="GI:1665773"
/translation="ATAGGSGADPGSRLLELLSFCVLLAGLCRNSVERKIYIPLN
KTAPCVRLNATHQIGCCSSISGDGVHVVEKEDELLOWLIDGNPNYVLLSKHF
ITDLMKLAGRTSRTAGLAVSLTKPSASGSPSVQCPNDGFGVYSNYPFEAFHRE

```

```

IOMNSLNGLAYEDFSPFIPILEDENETKVIKQYQDHNLSQNSAPTEPLCAMQLFS
HMAHIVISTATCMRRSSSTQSFISFVQLAAALQKADPTVTLPRNMFVFFQGETF
AATKDRSFFNVPAGQSAVASFVQLAAALQKADPTVTLPRNMFVFFQGETF
DYIGSSRMVYDEMEKGFVQLENSFVQLAAALQKADPTVTLPRNMFVFFQGETF
VEDLLATLEKSGAGVPAVILRRPNQSPQLPPSPQLRRNLSGVVLAHDSGAFHNK
YQSIYDTAENINVSYPLEWSPEDLNFTDTAKALADVATVLGRALYELAGTNTSD
TVQADPOTVTRLLYGLFLIKANNWFOSILRODLRSYLGDPLOHYIYAVSPNTYV
QYALANLTGVVNLTRQCCODPKVPSEKDLVEYSWVOGPLHNSHETDLRPRVSTA
RLALSPAFELSWSSSTESTWTSERWDIRAFILIASKELELITLVGVGGLIFS
LIVTYCINAKADVLFIAPREPVAVS*
BASE COUNT 648 a 765 c 646 g 746 t
ORIGIN
Query Match 95.1%; Score 2804.2; DB 91; Length 2805;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2803; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 145 ggctacggcaggggtgctctgggggtgacccgggaagtcggggtctccttcgcttct 204
|||||
Db 1 GGCTACGGCAGGGGTGCTCTGGGGCTGACCCGGGAAGTCGGGTCTCCTTCGCTCT 60
QY 205 gtctttcgtctcactagcaggtttgtgcaggggaaactcagtgaggaggaatata 264
|||||
Db 61 GTCTTCTGCGTCTCTAGCAGGTTTGTGCGGGGAACCTCAGTGGAGAGGAATATA 120
QY 265 tatcccttaataaacacgctccctgtgttcgctgctcaacgccactcactcagattg 324
|||||
Db 121 TATCCCTTAAATAAACACAGCTCCTGTGTTGCGCTGCTCAACGCCACTCATCAGATGG 180
QY 325 ctgcagctctcaattagtgagacacaggggtttatccactagtagagaaaggaggga 384
|||||
Db 181 CTGCCAGTCTTCAATTAGTGAGACACAGGGGTTTATCCACGTAGTAGAGAGAGAGGA 240
QY 385 cctacagtggttattgactgctgagcccaaccccccttaccatggttctgctggagagaa 444
Db 241 CCTACAGTGGTATTGACTGATGCGCCCAACCCCTTACATGTTCTGCTGGAGAGCAA 300
QY 445 gcattttaccaggatttaattggagaagctgaaaggaggagaaaccagccgaattgctgtct 504
Db 301 GCATTTTACCAGGATTTAATTGGAGAGCTGAAAGGGAGAACACAGCCGAATTGCTGTCT 360
QY 505 tgcagtgctcttgcagcagccagctcctgcctcaggttctctcctagtgacagtgccc 564
Db 361 TGCAGTGTCTTGACCAAGCCAGTCTGCTCAGGCTTCTCTCTAGTGTACAGTGGCC 420
QY 565 aaatgatgggttgggttttactccaattcctatggcgagagtttgcctcactcagaga 624
Db 421 AAATGATGGTGGTGGTGTACTTCAATTCCTATGCGCCAGAGTTTCTCACTCAGAGAGA 480
QY 625 aatacagtggaattcgtcggcaatggttggcttatgaagacttttagtttccccattt 684
Db 481 AATACAGTGGAAATTCGCTGGGCAATGGTTTGGCTTATGAAGACTTTAGTTTCCCATCTT 540
QY 685 tcttcttgaagatgaaatgaacaaagtcacatcaagcagtgcttatcaagatcacacct 744
Db 541 TCTTCTGAAGATGAAATGAACCAAGATCATCAGCAGTGTCTATCAAGATCAACACT 600
QY 745 gagtcgaatggtcagcaccacacttcccaactatgtgcaatgcagctcttttccacacat 804
Db 601 GAGTCAGAAATGGCTCAGCACCAACCTTCCCACTACTGTGCTCAGTGTGCTTTTTCACACAT 660
QY 805 gcagctgtcactcagcactgccactcagtcaggcgagctcctccatccaaagcaacttcag 864
Db 661 GCATGCTGTATCAGACTGCCACTGCTGATGGCGGAGCTGCCATCCAAAGCAGCTTCAG 720
QY 865 catcaacccagaaatcgtctgtgacccctctgtgattacaaatgtgtggagcatgctaaa 924
Db 721 CATCAACCCAGAAATCGTCTGTGACCCCTGTCTGATACAAATGTGTGGAGCATGCTAAA 780
QY 925 gcctataatacaactggagcattaaagcctcagacagaggttgtgtgtgctgcccaccc 984
Db 781 GCCTATAATAACAACCTGGGACATTAAAGCCTTGACGACAGAGGTTGTGTGTGCTGCCACCG 840

```



